

Rozšíření a využití webového portálu o mapových službách

Jan Jiránek¹

Katedra mapování a kartografie, FSv, ČVUT, Thákurova 7,
177 29, Praha 6, Česká republika,
jan.jiranek@fsv.cvut.cz

Abstrakt. Cílem diplomové práce, která volně navazuje na bakalářskou práci je návrh a realizace aktivního vyhledávání volně přístupných OGC Web Service (Web Map Service, Web Feature Service, Web Coverage Service) serverů. Tato aplikace pak bude integrální součástí portálu o webových geodatových službách. Na portálu bude existovat možnost prohlížení a vyhledávání metadat WMS serverů. Zároveň bude zřízena služba pro externí přístup (rozhraní) k aplikaci aktivního vyhledávání serverů. Portál může být využit státní správou, samosprávou, soukromým sektorem, jinými institucemi a osobami pracující s technologiemi webových geodatových služeb.

Klíčová slova: Web Map Service, Web Feature Service, Web Coverage Service, OGC Web Service, Open Geospatial Consortium, Google, Keyhole Markup Language, metadata, UMN Mapserver, mapové služby

Abstract. The aim of this Diploma Thesis, which is a continuation of my Bachelor work is a draft and an implementation of an active search of freely accessible OGC Web Service (Web Map Service, Web Feature Service, Web Coverage Service) servers. This application will be an integral part of the web portal concerning web geodata services. There will be a possibility to browse and search the metadata of WMS servers on the portal. A service for an external access (interface) to an application will be set up that enables active server search. The web portal can be used by the state administration, self-administration, private sector and by other institutions and subjects working with web geodata services.

Keywords: Web Map Service, Web Feature Service, Web Coverage Service, OGC Web Service, Open Geospatial Consortium, Google, Keyhole Markup Language, metadata, UMN Mapserver, map service

1 Úvod

Standardy WMS, WFS a WCS tvoří páteřní kostru webových geodatových služeb OGC (OWS) a naplňují tak předpoklady interoperability v GIS. Vedle

nich vznikají nebo již existují další standardy, které mají rozšířit a doplnit koncept OGC Web Service (OWS). Jedná se například o katalogové služby (OGC CAT).

Mojí snahou je navázat na poznatky a znalosti získané při psaní bakalářské práce na téma **Webový portál o mapových službách**, případně je upravit nebo doplnit o nová zjištění. Prvotním cílem je vytvoření aplikace aktivního vyhledávání volně přístupných OWS serverů, která bude schopna zpracovat zjištěné WMS servery a jejich metadata do databáze. Pro přístup k těmto datům bude vytvořeno webové rozhraní, kde si uživatelé budou moci prohlížet či vyhledat OWS servery dle definovaných metod. Pro případného zájemce seznamu aktivních OWS serverů bude existovat možnost exportu URL serverů.

Neméně významnou snahou je pokusit se navrhnout externí službu pro přístup k vyhledávání v databázi metadat OWS serverů pomocí rozhraní HTTP metody GET (KVP). Služba by tak mohla sloužit jakýmkoliv aplikacím, které chtějí vyhledat OWS servery, resp. jejich data, pomocí různých metod a následovně si je připojit do klienta.

Výstupů této práce bude vícero. Nejdůležitějším je webový portál o mapových službách, který se bude skládat z informativní sekce o technologii webových geodatových služeb a databázové sekce obsluhující aplikaci aktivního vyhledávání a správy metadat OWS serverů. Takovýto portál by měl všestranné využití pro kteréhokoliv zájemce o tuto technologii.

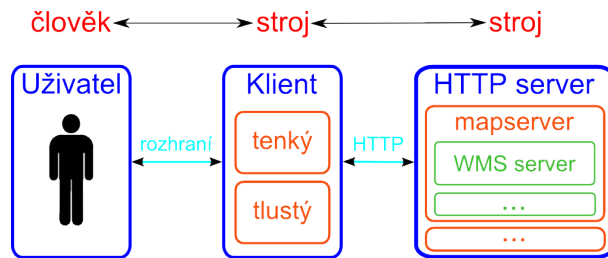
2 Úvod do distribuce geodat

Webové geodatové služby jsou přínosným rozšířením tradičních GISů. Umožňují distribuovat geodata v různých podobách (transformacích) prostřednictvím internetu (intranetu). Dávají tak rychlou a snadnou možnost standardizované publikace těchto dat především pro správce dat. Tyto služby lze pak začlenit do nejrozličnějších softwarových řešení či jiných standardů. Příkladem může být využití služeb webovými mapovými aplikacemi nebo integrací s dalšími technologiemi, popř. standardy, jakým je průlet nad DMT s pokryvem z WMS vrstvy. Tato aplikace navíc může pracovat v okně webového prohlížeče v objektu typu Shockwave.

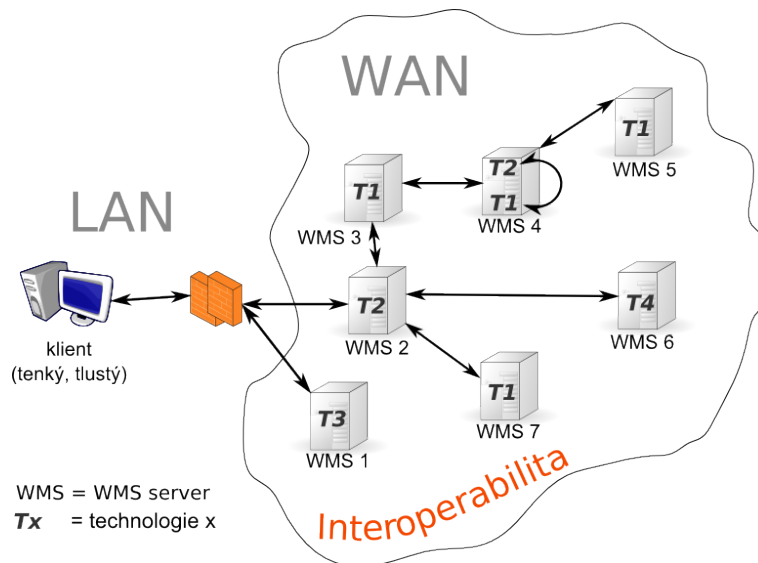
Pokud jsou tyto služby standardizovány mezinárodní neziskovou organizací Open Geospatial Consortium, Inc.® (OGC), můžeme hovořit o naplňování předpokladu interoperability v GIS. Mezi základní pilíře webových geodatových služeb OGC patří Web Map Service (WMS), Web Feature Service (WFS) a Web Coverage Service (WCS).

2.1 Architektura server-klient

Technologie publikace geodat prostřednictvím internetu (intranetu) je založena na vzájemné interakci mapového serveru, klienta a uživatele (člověka) názorněji vysvětleno na obr. č. 1.



Obr. 1. Schéma komunikace mezi mapovým serverem, klientem a člověkem (zdroj: vlastní zpracování)



Obr. 2. Ukázka interoperability v GIS - distribuce WMS při různých technologických řešeních tzv. kaskádování (zdroj: vlastní zpracování)

Schopnost porozumění mezi klientem a mapovým server je umožněna díky existenci interoperability v prostředí GIS. Zjednodušeně řečeno se jedná o to, že dvě rozdílné technologie mluví stejným jazykem a rozumí významu jednotlivých vět (obr. č.2).

Tato situace na počátku vzniků GIS chyběla. Často si jednotliví tvůrci systému vytvářeli vlastní uzavřené geodatové formáty, webové služby a jiné související produkty. V současnosti jsou tyto bariéry díky OGC odstraňovány.

3 Způsoby vyhledávání OGC Web Service

S poměrně rychlým nástupem technologie webových geodatových služeb OGC začaly vznikat první volně přístupné servery poskytující přístup ke geodatům. S růstem počtu těchto serverů se rozšiřovala nabídka software, který uměl tyto služby publikovat. Z počátku vznikaly především WMS servery, které publikují geodata v té nejméně zneužitelné formě z pohledu správce dat. S růstem počtu volně přístupných WMS serverů se pak často v odborných člancích a na webových stránkách objevovaly seznamy těchto serverů.

Představa, že lze k internetu přistupovat tak jako k webovým geodatovým službám, je patrně deprimující. Doby, kdy uživatel přistupoval k internetu s notýskem zajímavých adres, jsou pryč. Vznikly internetové vyhledávací portály (např. Google), které ví o těchto webových stránkách a aktivně vyhledávají další, aby mohly uživateli nabídnout seznam stránek výsledků jeho vyhledávání. Tento koncept pomalu nastupuje i do oblasti OWS. OGC k tomu vyvíjí a inovuje standard Catalogue Services for Web (CSW), označovaný někdy jako katalogová služba (CAT).

3.1 „WMS seznamy“

Na internetu a v odborných člancích se běžně vyskytují výčty mapových služeb. Pro jednoduchost je nazvěme „WMS seznamy“, ačkoliv neodkazují jen na WMS servery, ale jejich převaha je pro ostatní služby zdrcující.

Starší verze těchto výčtů pak obsahují název serveru a jeho URL. Pokud se jedná o propracovanější variantu, obsahuje i seznam dostupných vrstev a jejich vlastností. Posuneme-li se v čase, začaly se objevovat weby umožňující vyhledávání mezi nimi [5]. Největším záporem, který může za brzký zánik významu těchto seznamů, je jeho nesoulad s aktuálním stavem.

3.2 Globální internetové vyhledávací portály

Kdokoliv, kdo dnes pracuje s internetem, tak využívá v určité míře webové vyhledávací portály (Google, Yahoo, Seznam, Centrum, aj.). Jsou nyní nedílnou součástí infrastruktury WWW.

Portál Google je v současnosti nejpoužívanější světový internetový vyhledávač, který každodenně obslouží přes 400 milionů dotazů a indexuje 25 bilión webových stránek . Na Google probíhá 53,6% vyhledávání¹ uskutečněných na světě.

Běžný uživatel Google napíše vhodná klíčová slova do vyhledávacího pole a začne procházet výsledky seřazené dle relevantnosti. Pokud hledal specifitější výsledek a ve výsledcích vyhledávání ho nenašel, tak se pravděpodobně dopustil jedné z mnoha chyb při hledání na vyhledávači. Jednou z možností je, že blíže nespecifikoval dle klíčového slova svůj požadavek, nebo naopak nepoužil vhodné operátory upřesňující metody hledání.

Základními operátory jsou nejčastěji znaky „+“ (logické AND) a „-“ (logické OR). V případě hledání sousloví stačí dát výraz do „uvozovek“. Pokročilejšími možnostmi jsou např. operátory:

- **filetype:** Hledání souboru určitého typu, např. `filetype:PDF`.
- **site:** Hledání omezené jen na určitou doménu, např. `site:www.cvut.cz` **ISIC**.
- **inurl:** Hledání soustředěné jen na URL řetězec, např. `inurl:cvut`.

Při prohlížení stránek výsledků vyhledávání má Google všechny potřebné parametry aktuální stránky vloženy do párů parametr a hodnota parametru (KVP) metody GET protokolu HTTP. Tvar URL hledaného výrazu *ČVUT* vypadá takto:

```
http://www.google.com/search?hl=en&lr=&pwst=1&q=CVUT&start=20
```

Parametry ovlivňující hledání jsou pak:

- **hl:** Národní lokalizace daná kódem např. `en` (English), `cs` (Czech), ...
- **q:** Řetězec s parametry hledání (klíčová slova, operátory, aj.)
- **start:** ID prvního výsledku na stránce. Rozsah výsledků $< ID, ID + 10$).

3.3 Metadatové portály

Metadatový portál je webová aplikace poskytující přístup k spravovaným metadatům geodat pomocí rozhraní (služby). S tímto přístupem k publikaci metadat se počítá při budování INfrastructure for SPatial InfoRmation in Europe.

MICKA je webová aplikace společnosti Help Service - Remote Sensing s.r.o., která je jedním z nejzkušenějších odborníků v oblasti metadat geodat v ČR a ve využívání Open Source projektů ve svých produktech. MICKA (metainformační katalog) je nejnásazovanější metainformační portál v oblasti státní správy a samosprávy. Mezi nejvýznamnější realizace patří **Metadatový portál životního prostředí** vytvořený ve spolupráci s agenturou CENIA. Aplikace² podporuje OGC CSW 2.0.2 (katalogová služba), standardy ISO 19115 / 19119 / 19139 (metadata prostorových dat a služeb), ISO 15836 (metadata Dublin Core), metadatový profil INSPIRE, import metadat z jiných profilů (ESRI ArcCatalog, aj.), RSS kanály pro zobrazování změn a jiné funkce.

¹ K srpnu 2007.

² Vzataženo k listopadu 2008.

INSPIRE je iniciativou Evropské komise pro vytvoření stejnojmenné směrnice Evropské komise a Rady. Za cíl si klade vytvořit evropský legislativní rámec potřebný k vybudování evropské infrastruktury prostorových informací [6].

Směrnice INSPIRE znamená pro členské země EU v rozmezí let 2007-2009 začlenění směrnice do národních legislativ včetně prováděcích vyhlášek. Ministerstvo životního prostředí plánuje zřídit Národní geoportál INSPIRE, který bude široké veřejnosti zpřístupňovat jmenovaná prostorová data. Metadaty (podle implementačních pravidel INSPIRE) budou opatřena všechna data i služby na geoportálu.

Testovací metadatový portál INSPIRE [8] je již nějakou dobu v provozu. Umožňuje efektivně vyhledávat mezi geodaty. Bohužel zatím³ se řeší ještě některé problémy, jako např. podpora vícejazyčného metadatového dokumentu.

3.4 Návrh způsobů vyhledávání OWS serverů

Technologie katalogových služeb (CAT) dokáže přistupovat, publikovat a vyhledávat mezi metadaty, které jsou publikovány dle možností, s nimiž katalogová služba umí nakládat. Fakticky to znamená, že katalogová služba ví o datech (službách), jenž jsou publikovány tak, aby služba s nimi uměla nakládat (řetězení katalogových služeb).

V případě, že chceme nalézt službu, např. WMS, která není publikována pomocí katalogových služeb, náš přístup selže. Pro vyhledání těchto služeb se musí provést aktivní operace (kroky).

Implementace znalostí technologie OWS: Máme-li URL a zatím nevíme, zda je adresou k jedné z webových služeb OWS, můžeme vytvořit předpoklad existence služby z rozboru URL jako řetězce. Pro OWS (WMS, WFS, WCS) musí být URL protokolu HTTP popř. HTTPS. Ze znalosti mapových serverů distribuujících geodata pomocí OWS, lze odvodit následující předpoklady:

- URL query obsahuje hodnotu a parametr `SERVICE=WMS`, `SERVICE=WFS` nebo `SERVICE=WCS`
- URL query obsahuje parametr `map` s cestou k souboru Mapfile (`.map`). Příklad `?map=/data/foo.map` (UMN Mapserver - WMS, WFS a WCS).
- URL adresa dokumentu obsahuje další řetězce např. `mapserv`, `mapserv.exe`, `com.esri.wms.Esrimap`, `com.esri.wfs.Esrimap`, `geoserver`, `deegree`, ...

Každý z mapových serverů má své specifikum, které se často odráží ve tvaru URL. Stačí jen na něj reagovat.

4 Aplikace vyhledávání OGC Web Service

Vytvořená aplikace **OWS active search** je schopna dle definovaných metod aktivně vyhledávat volně přístupné OGC Web Service v prostoru internetu.

³ Prosinec 2008.

Díky tomuto přístupu může nalézt i nové veřejně nepublikované OWS servery. Nejčastěji se pak může jednat o tenkého klienta vytvořeného pomocí UMN Mapserveru, který publikuje geodata jen prostřednictvím webového uživatelského rozhraní.

Aplikace pak dokonce dokáže rozeznat, pod jakými standardy OWS mapový server funguje. Pokud by se u určitého URL tvrdilo, že poskytuje přístup k datům jen prostřednictvím WMS, tak aplikace je to schopna ověřit, popř. opravit tento předpoklad.

4.1 Důležité znalosti technologie OWS

Pro vytvoření aplikace aktivního získávání OWS serverů a správy metadat je zapotřebí znát specifika technologie OWS [4], [2], [1]. Příkladem může být tag `SCALEHINT` (WMS < 1.1.1) × `SCALE` (WMS 1.3.0) vyjadřující měřítkový rozsah (dostupnost) WMS vrstev.

Prvotním úkonem, jenž vykoná klient při komunikaci s OWS (WMS, WFS nebo WCS) serverem, je zjištění všech možností serveru. Jedná se o sestavení validního dotazu `GetCapabilities` (dále jen GC dotaz) klientem a zpřístupnění metadat ve formátu XML OWS serverem. Postup jak sestavit validní dotaz GC dotaz je jednoduchý, ale na druhou stranu nabízí více možností, které je potřeba ověřit (22 možností). Je to z důvodu odchýlení od standardu a z důvodu existence vícero vývojových verzí jednotlivých služeb.

4.2 Metody vyhledávání pravděpodobných URL OWS serverů

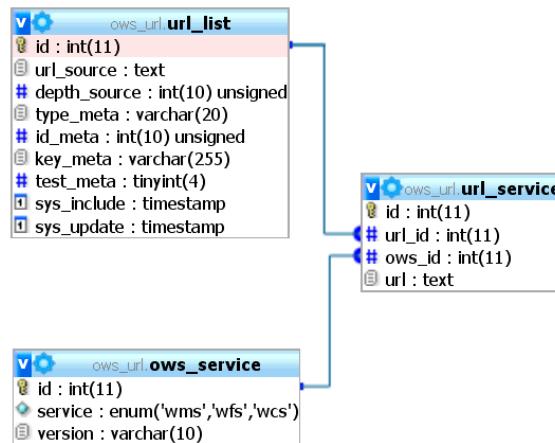
Rozborem a znalostí technologie publikace dat pomocí webových geodatových služeb OGC (OWS) a jednotlivých softwarových řešení lze vytvořit odlišný přístup aktivního vyhledávání OWS serverů. V předchozí kapitole je popsán návrh na nový způsob vyhledávání OGC Web Service serverů, např. pomocí specifického řetězce znaků vyskytujícího se v URL OWS serveru určité technologie. Nástrojem pro takovéto vyhledávání jsou webové vyhledávací portály, které poskytují přístup k výsledkům vyhledávání prostřednictvím API.

Tento postup vyžaduje složení specifického URL serveru GOOGLE (dotaz na server) pro zpřístupnění stránky vyhledávání, ze které se pak získají URL **výsledků vyhledávání** (pravděpodobné OWS - dle klíčových slov).

Vytvořenou metodou hledání OWS serverů je již zmíněný způsob prostřednictvím Google vyhledávání. Další vhodnou metodou je přístup využívající tzv. „WMS seznamy“. Jedná se o definované hledání na konkrétních portálech, např. [5].

4.3 Popis aplikace

Jádrem vlastní aplikace pro aktivní vyhledání a ověřování serverů OGC Web Service (OWS) jsou vytvořené PHP třídy z hlediska objektivě orientovaného programování (OOP). Každá z tříd má svůj specifický účel a význam.



Obr. 3. ER diagram aplikace pro aktivní vyhledávání OWS služeb (zdroj: vlastní zpracování)

Třídy poskytují jednotlivým skriptům potřebné metody. Každý skript pak vykonává určitou činnost, která je nezbytná pro správný chod aplikace. Pro pravidelné spouštění skriptů slouží systém CRON. Soubory pravidel pro hledání OWS serverů jsou uloženy v XML (konfigurační soubor).

Pro chod aplikace (datové úložiště) byla využita databáze MySQL. Aplikace, která získává pravděpodobné URL OWS serverů a analyzuje je na službu OWS a její verzi, se skládá ze tří tabulek (obr. č. 3). Celá aplikace se skládá přibližně z 20 tabulek sloužících především pro ukládání metadat.

4.4 Zhodnocení úspěšnosti vyhledávání OWS

Aplikace byla testována na mém osobním notebooku⁴. V době testování aplikace na hledání OWS dle vybraných metod docházelo k jejímu ladění a optimalizaci. Pro vyhledání odkazů bylo zapotřebí nechat aplikaci soustavně pracovat asi 1 měsíc. Při této činnosti bylo získáno asi přes 1 milión URL.

Metoda hledání na tzv. „WMS seznamech“: Výsledky vyhledávání zaměřené na tzv. „WMS seznamy“ jsou uvedené v tabulce č. 2. Jak je z tabulky patrné, tak převážná část OWS byla získána z prvotního URL (hloubka 0). Dále lze dodat, že pokud tyto seznamy nejsou udržovány v souladu se skutečností, tak velice rychle zastarávají.

⁴ Fujitsu Siemens Amilo Pro V3545, Windows XP Professional SP2, Intel Core 2 Duo 1660 MHz, DDR2 2048 MB

Metoda hledání pomocí „Google“: Výsledky vyhledávání OWS pomocí „Google“ jsou vůči předešlé metodě už z principu jiné. Výsledky jsou uvedeny v tabulce č. 1. Při porovnání jednotlivých metod hledání lze usoudit, že operátor `inurl:` a znalost technologie mapových serverů umožňují poměrně úspěšné nalezení OWS serveru.

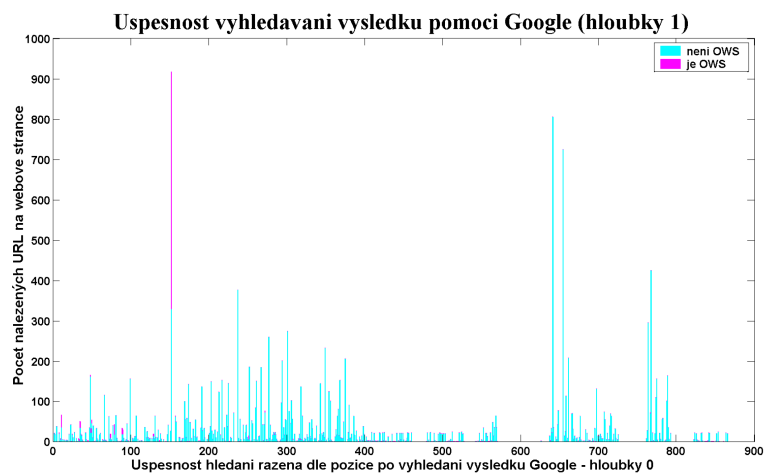
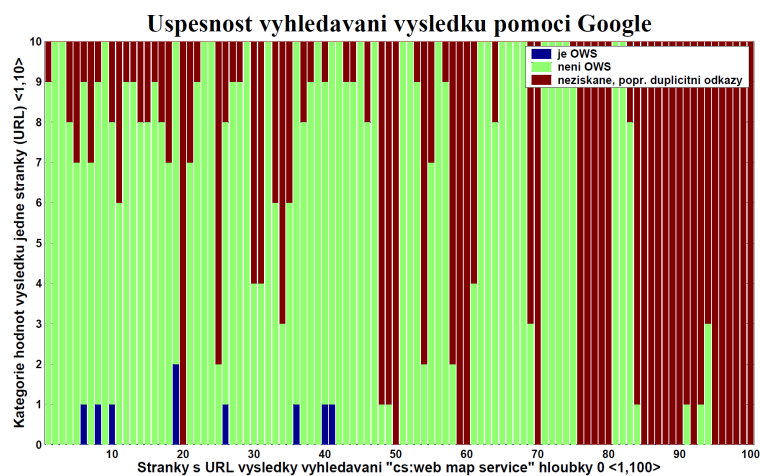
Nejdéle běžící proces hledání probíhal na klíčové slovo „web map service“. Hledání pak trvalo soustavně přes týden. Čas potřebný na práci je pouze orientační, protože na tomto odkazu se provádělo první ladění aplikace. Nejvíce funkčních OWS URL bylo nalezeno v hloubce vnoření 1. Úspěšnost hledání pak dokazují obrázky grafů č. 4.

Tabulka 1. Přehled úspěšnosti vyhledávání OWS pomocí Google a vlastního robota

ID Výraz	Hloubka vyhledávání					
	0			1		
	je	není	[%]	je	není	[%]
1 web map service	11	714	1.5	781	16 483	4.5
2 inurl:esrimap	115	188	38.0	0	579	0.0
3 inurl:wms	12	643	1.8	162	17 047	0.9
4 inurl:wfs	32	519	5.8	30	6 643	0.4
5 inurl:wcs	4	524	0.8	6	9 184	0.1
6 inurl:mapserv	259	383	40.3	3	698	0.4
inurl:service=wms						
7 +OR+	353	119	74.8	3	945	0.3
inurl:service=wfs						
8 inurl:service=wcs	29	22	56.9	1	191	0.5

Tabulka 2. Přehled úspěšnosti vyhledávání OWS pomocí vlastních metod vyhledávání na „WMS seznámech“

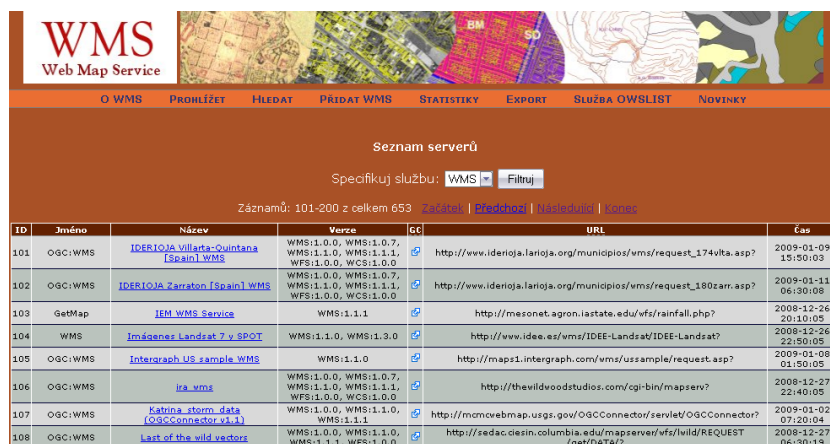
ID Web	Hloubka vyhledávání								
	0			1			2		
	je	není	[%]	je	není	[%]	je	není	[%]
1 ogcservices	776	241	76.3	1	7	12.5	0	29	0.0
2 skylab	619	376	62.2	6	21	22.2	7	180	3.7
3 wms-sites	489	57	89.6	0	2	0.0	0	11	0.0
4 GIDB	567	0	100.0	0	0	0.0	0	0	0.0



Obr. 4. Úspěšnost vyhledávání OGC Web Service pomocí vlastní aplikace na Google (hloubka 0, 1) dle klíčového slova „web map service“ (zdroj: vlastní zpracování)

4.5 Chyby „GIS specialistů“

Aplikace vyhledávání aktivních volně přístupných OWS (WMS, WFS a WCS) serverů odhalila zajímavou skutečnost. Někteří zřizovatelé mapových serverů nemají často potuchu, jak daná aplikace pracuje a publikují pomocí ní svá cenná geodata určená jen pro konkrétní použití, jako např. webová aplikace (tenký klient). Např. UMN mapserver dokáže implicitně komunikovat na WMS standardu, aniž by někdo službu povoloval.



ID	Jméno	Název	Verze	IC	URL	čas
101	OGC:WMS	IDERIOJA Villarta-Cuintana [Spain] WMS	WMS:1.0.0, WMS:1.0.7, WMS:1.1.0, WMS:1.1.1, WFS:1.0.0, WCS:1.0.0	✓	http://www.iderioja.larioja.org/municipios/wms/request_174vita.asp?	2009-01-09 13:50:03
102	OGC:WMS	IDERIOJA Zaraton [Spain] WMS	WMS:1.0.0, WMS:1.0.7, WMS:1.1.0, WMS:1.1.1, WFS:1.0.0, WCS:1.0.0	✓	http://www.iderioja.larioja.org/municipios/wms/request_180zarr.asp?	2009-01-11 06:30:08
103	GetMap	ICM WMS Service	WMS:1.1.1	✓	http://mesonet.agron.iastate.edu/wfs/rainfall.php?	2008-12-26 20:10:05
104	WMS	Imágenes Landsat 7 y SPOT	WMS:1.1.0, WMS:1.3.0	✓	http://www.idee.es/wms/IDEE-Landsat/IDEE-Landsat?	2008-12-26 23:50:05
105	OGC:WMS	Intergraph US sample WMS	WMS:1.1.0	✓	http://maps1.intergraph.com/wms/ussample/request.asp?	2009-01-08 01:50:05
106	OGC:WMS	ira wms	WMS:1.0.0, WMS:1.0.7, WMS:1.1.0, WMS:1.1.1, WFS:1.0.0, WCS:1.0.0	✓	http://thewildwoodstudios.com/cgi-bin/mapserv?	2008-12-27 22:40:05
107	OGC:WMS	Kabina storm data (OGCconnector v1.1)	WMS:1.0.0, WMS:1.1.0, WMS:1.1.1	✓	http://mcmceebmap.usgs.gov/OGCConnector/servlet/OGCConnector?	2009-01-02 07:20:04
108	OGC:WMS	Last of the wild vectors	WMS:1.0.0, WMS:1.1.0, WMS:1.1.1, WFS:1.0.0	✓	http://sedac.ciesin.columbia.edu/mapserver/wfs/wild/REQUEST/GET/DATA/?	2008-12-27 06:30:19

Obr. 5. Ukázka nové verze prohlížení metadat OWS serverů „Webového portálu o mapových službách“ (zdroj: vlastní zpracování)

4.6 Úprava „Webového portálu o mapových službách“

Webový portál o mapových službách, jenž byl výsledkem mé bakalářské práce, umožňoval přístup k informacím o technologii WMS a souvisejících heslech. Mým pokusem dále bylo publikovat metadata WMS serverů. Principem a zároveň jak se ukázalo problémem tohoto způsobu bylo to, že jsem očekával ze strany návštěvníků aktivitu ve smyslu, že doplní nové WMS servery. Tímto přístupem tam nepřibyl skoro žádný server. Tento nový přístup je založen na aktivním vyhledávání základní palety OWS (WMS, WFS, WCS) pomocí vytvořených metod, využívá znalostí příznaků specifických typů mapových serverů.

Sekce informačních článků byla rozšířena o zajímavá a aktuální témata (ČÚZK WMS server, aj.). Celá databázová sekce publikující a spravující metadata OWS, resp. WMS serverů, byla kompletně předělaná. Jedná se o upravení stránek, tzn. Statistika, Prohlížet, Hledat, Přidat WMS, Export a Novinky. Novým prvkem je stránka **Služba OWSLIST**, jejíž účelem je zřízení externího přístupu k vyhledávání.

5 Jednoduchá aplikace (KVP) pro přístup k seznamu OWS

Představte si, že se na jednom místě nachází určitý seznam dat, jehož smysl je, že má být dále publikován. Na tomto místě je zapotřebí zřídit službu, která pomocí definovaného rozhraní umožní přístup k těmto datům.

Ze strany OGC je v tomto duchu pro geodata vyvíjen a inovován standard **Catalog Services for the Web (CSW)** [4], jehož poslední vývojová verze 2.0.2⁵ ze dne 23. 2. 2007 je zatím ve fázi druhých oprav⁶.

Cílem bylo vytvořit jednoduchý návrh služby pro přístup k spravovaným metadatům **OWSLIST**, s dodržením konceptu OWS především z hlediska přístupu k tvoření dotazů na server. V žádném případě není mou snahou konkurovat katalogovým službám OGC, které jsou dlouhodobě vyvíjeny a důmyslně propracovány. Důkazem toho je skutečnost, že poslední referenční dokument CSW 2.0.2 má 218 stran, dále odkazuje na jiné dokumenty, jež k tomuto standardu nedílně patří, např. dokumenty metadatových profilů. Jak je patrné, tak CSW někdy označována jako CAT (katalogová služba), je daleko rozsáhlejší a náročnější na implementaci než samotné webové geodatové služby OGC (WMS, WFS a WCS).

5.1 Popis služby OWSLIST pro přístup k spravovaným metadatům OWS serverů

Služba **OWSLIST** je pokusem o návrh a zároveň průzkum možností zřízení služby pro přístup k metadatům OGC Web Service, resp. seznamu OWS serverů (WMS, WFS a WCS).

Takováto aplikace je navržena v rámci konceptu OWS, kdy dotazy na server jsou kladeny pomocí KVP obsaženého u URL dotazu GET metody HTTP(S). Inspirací konceptu (logiky) dotazů metody GET protokolu HTTP byla specifikace OGC WMS 1.3.0 (tvar URL, rezervované znaky hodnot parametru query URL, aj.). Společnými parametry pro dotaz na službu **OWSLIST** jsou **SERVICE**, **VERSION**, **REQUEST**.

Řazení výsledků dotazů: Pro hledání v metadatech při úspěšném výskytu větším jak jeden nález je vhodné zavést řazení výsledků vyhledávání. Jednalo by se o období řazení výsledků **ORDERBY** při výběru dat z databáze. Daná data jsou natolik specifická, že je potřeba teoreticky nadefinovat možnosti jejich řazení.

Jednou z možností určení relevantnosti při hledání dle bodu by byla poloha tohoto bodu vůči středu boxů. V případě hledání dle boxu (např. ČR, okres, ORP, Kú, aj.), lze jednoduše určit každému boxu v relaci relevanci vycházející z překryvu ploch (0 %, 0 % < x < 100 %, 100 %).

Pro fulltextové hledání by bylo vhodné zavést váhy pro atributy kategorií metadat (server, vrstva nebo stylu vrstvy) WMS serverů, obsažené v XML

⁵ Referenční dokument OGC 07-006r1.

⁶ angl. Corrigendum 2 Release

dotazu `GetCapabilities`. Prostým váženým průměrem u každého záznamu (server, vrstva, styl vrstvy) by se vypočetl váhový koeficient vyhledávání.

$$r_{fulltext} = \frac{\sum_{i=1}^n (V_i P_i)}{\sum_{i=1}^n P_i} \quad (1)$$

V_i počet výskytů hledaného slova v atributu i

P_i váha atributu i

$r_{fulltext}$ relativnost daného záznamu

Příklady subjektivních vah pro tagy fulltextového vyhledávání kategorie WMS vrstva: <Name> - 1, <Title> - 2, <Keyword> - 3, aj.

Všechna tato řešení, se již v různých podobách začínají objevovat na metadataových vyhledávacích portálech [8], [7].

Tabulka 3. Parametry URL dotazu `GetList` služby `OWSLIST`

Parametr	P/V	Význam
<code>SERVICE=OWSLIST</code>	P	identifikace služby
<code>VERSION=0.0.1</code>	P	identifikace verze služby
<code>REQUEST=GetList</code>	P	název dotazu
<code>OUTPUTFORMAT=output_format</code>	P	formát odpovědi služby po dotazu (MIME)
<code>OWS=ows_list</code>	V	identifikace služby (služeb)
<code>LEVEL=level</code>	V	identifikace služby
<code>METHOD=method_list</code>	P	metoda vyhledávání OWS služeb
<code>GEOQ=query_value</code>	P*	parametry geografického hledání
<code>FULLTEXTQ=query_value</code>	P*	parametry fulltextového hledání
<code>OUTPUTSCHEMA=schema_value</code>	V	druh výstupního schématu
<code>USERSHEMA=schema_list</code>	V	jmenování výstupních parametrů
<code>ORDERBY=orderby_list</code>	V	kategorie priority řazení
<code>EXCEPTIONS=output_format</code>	V	formát chybové hlášky na službu
<code>OPERATOR=operator</code>	P*	operátor relace metody <code>geo</code> a <code>fulltext</code>

Poznámky:

– Význam hodnot atributu „přítomnost“ *P* - povinný, *V* - volitelný, * - s podmínkou.

Dotaz GetList na webovou službu OWSLIST zpřístupňuje v definovaných formátech přístup k databázi OGC Web Service (OWS). Současně je to jediný povolený dotaz na aplikaci OWSLIST REQUEST=GetList.

Žádost na server lze specifikovat na požadovanou kategorii (subkategorii) dat, např. WMS (server, layer, style), WFS (server, layer), aj. Pro selekci specifických dat (hledání) existují dvě metody, které se dají navzájem kombinovat s využitím logického AND a OR. Fulltextové vyhledávání, jenž je dobře známé z webových vyhledávacích portálů je první z možností. Další možnost je geografické vyhledávání dle bodu (boxu) v souřadnicích WGS84.

V případě, že dotaz je složen nekorektně tak server odpoví chybovou hláškou (TXT, XML). Klient může požádat o data uložená v databázi ve specifikovaných formátech (XML, CSV). Pro ovlivnění úrovně podrobností výsledků dotazů existuje parametr OUTPUTSCHEMA. Podrobněji je tato část probrána v [1].

Výpis 1.1 Příklad dotazu GETLIST služby OWSLIST: fulltextové a geografické hledání dle boxu; řazeno dle fulltextové relevance a bbox relevance

```
http://www.example.com/service.php?SERVICE=OWSLIST&
REQUEST=GETLIST&VERSION=0.0.1&METHOD=fulltext,geo&
GEOQ=12.33,45.76,17.88,50.77&OUTPUTFORMAT=text/xml
FULLTEXTQ=soil&ORDERBY=fulltext,bbox&OPERATOR=AND
```

6 Závěr

Cílem diplomové práce bylo rozebrat možnosti a způsoby vyhledávání aktivních volně přístupných OGC Web Service (OWS) serverů. Následně jsem tato zjištění využil při realizaci aplikace aktivního vyhledávání OWS serverů („robot“), jenž je nyní součástí webového portálu [3].

Mým zájmem bylo také prozkoumat možnosti vizualizace metadat geodat, která jsou svou podstatou opět geodata. Zároveň mě zajímalo, jakým způsobem by šlo integrovat OWS (WMS, WFS a WCS) do software Google Earth. Další mojí snahou bylo zjistit způsoby přechodu mezi těmito službami (GIS-ETL). Více o tomto tématu se dozvíte v [1].

Největší úspěch spatřuji v tom, že mnou vytvořené aplikaci aktivního vyhledávání OWS (WMS, WFS, WCS) serverů, se podařilo nalézt 1 614⁷ aktivních WMS serverů a 358 452 jejich vrstev. Na území ČR se nachází 88 aktivních WMS severů (BBOX, .cz). Je potřeba si uvědomit, že aplikace k tomu musela fakticky navštívit asi přes 1 milión URL a ty otestovat, zda jsou OWS službou. Dané servery dokáží často pracovat na vícero vývojových verzí standardu WMS popř.

⁷ Nalezeno na soukromém notebooku, přibližně 1. měsíc aktivního vyhledání. Na serveru GEO3 vytvořené webové aplikace probíhá opětovné vyhledání z tohoto seznamu, tudíž se počet nalezených serverů nemusí shodovat se stavem aplikace.

WFS a WCS. Pro takovýto počet serverů je podíl WMS 1.0.0 57 %, WMS 1.0.7 36 %, WMS 1.1.0 94 %, WMS 1.1.1 98 % a WMS 1.3.0 5 %.

Aplikace aktivního vyhledávání pak dokázala nalézt i WMS servery, jenž nikde nejsou na internetu publikovány (zmíněny), ba dokonce servery, u nichž vlastník, nepředpokládal jejich publikaci.

Dalším výstupem z mé diplomové práce bylo zřízení (oprava) článků české mutace internetové encyklopedie **Wikipedia** s otevřeným obsahem, na jejíž tvorbě spolupracují dobrovolní přispěvatelé z celého světa. Jedná se především o články Web Map Service (WMS), Web Feature Service (WFS), Keyhole Markup Language (KML) a European Petroleum Survey Group (EPSG).

Portál je zprovozněn na serveru GEO3 (*PHP 5.2.6, MySQL 5.0.51a*)

<http://geo3.fsv.cvut.cz/wms>,

katedry Mapování a kartografie, Fakulty stavební, ČVUT.

Možností, jak tento portál rozšířit (využít), je nepřeberné množství. Webové geodatové služby, resp. OGC Web Service (OWS), budují kolem sebe architekturu, která dává OGC geodatovým službám nové uplatnění či význam. Z několika nápadů rozvoje portálu uvádím např.:

- publikace spravovaných metadat dle standardu OGC Catalogue Service for Web (CSW)
- export metadat OWS serverů do XML profilů metadat ESRI, ISVS, Dublin Core, INSPIRE, aj.
- vizualizace metadat pomocí Google Earth nebo UMN MapServeru
- export vrstvy do souboru připojení, např. KML programu Google Earth nebo XWMS CAD software Microstation 8.11s
- vytvoření parseru na datové vrstvy pro jejich členění do vytvořeného katalogu kategorií a podkategorií (pro různé jazyky využití překladových služeb např. Google)
- instalace nového redakčního systému

Doufám, že vytvořená diplomová práce, resp. „Webový portál o mapových službách“ přispějí k tématu aktuální problematiky webových geodatových služeb. Rovněž vidím příležitost ve využití nově vybudované databázové části portálů, která by mohla případným zájemcům zpřístupnit pro ně tolik potřebná metadata ve formě URL WMS serveru.

Reference

1. JIRÁNEK, Jan. *Diplomová práce: Rozšíření a využití webového portálu o mapových službách*. Plzeň, 2009. 134 s. ČVUT. Vedoucí diplomové práce Petr Soukup. Dostupný z WWW: <<http://geo2.fsv.cvut.cz/~soukup/dip/jirane/jirane.pdf>>.
2. JIRÁNEK, Jan. *Bakalářská práce: Webový portál o mapových službách*. Plzeň, 2007. 83 s. ČVUT. Vedoucí bakalářské práce Petr Soukup. Dostupný z WWW: <<http://geo2.fsv.cvut.cz/~soukup/bkl/jirane/jirane.pdf>>.

3. JIRÁNEK, Jan. *Webový portál o mapových službách*. c2007-2008 [cit. 2009-05-04]. Dostupný z WWW: <<http://geo3.fsv.cvut.cz/wms/>>.
4. Open Geospatial Consortium, Inc ®. *OGC Homepage* [online]. c1994-2008 [cit. 2009-05-04]. Dostupný z WWW: <<http://www.opengeospatial.org>>.
5. *WMS Sites* [online]. [2006-2007] [cit. 2009-05-04]. Dostupný z WWW: <<http://wms-sites.com/>>.
6. CENIA. *INSPIRE - Infrastructure for SPatial InfoRmation in Europe* [online]. c2008 [cit. 2009-05-04]. Dostupný z WWW: <<http://www.czinspire.cz>>.
7. Ministerstvo životního prostředí. *Metadatový portál*. [online]. [2007] [cit. 2009-05-04]. Dostupný z WWW: <<http://mis.cenia.cz>>.
8. EC JRC. *Metadata Catalogue start page*. [online]. [2008] [cit. 2009-05-04]. Dostupný z WWW: <<http://www.inspire-geoportal.eu/catalogues.htm>>.