



Otevírání geografických dat

OpenGeoLabs s.r.o.

leden 2015

1	Úvod	7
1.1	Otevřená data	8
1.2	Uživatelé otevřených geodat	15
2	Technické a právní aspekty otevírání datových sad	19
2.1	Výběr licence pro otevřená geodata	19
2.2	Otevřené formáty, webové služby a distribuce otevřených geodat	19
2.3	Verzování dat a časové řady	27
2.4	3D Data	29
2.5	Metadata	29
2.6	Souřadnicové systémy	30
2.7	Závěrečná doporučení	31
3	Otevírání geografických dat	33
3.1	Doporučený postup otevírání geografických dat	33
3.2	Technické řešení infrastruktury otevřených dat	34
A	Příloha A: Open Source software pro jednotlivé úlohy	37
A.1	Geodatabáze PostGIS	37
A.2	Mapový server MapServer	38
A.3	Mapový server GeoServer	38
A.4	Knihovna GDAL	38
A.5	PyCWS	38
A.6	GeoNetwork	38
B	Příloha B: OGC WMS – dimenze a časové řady	39
B.1	Příklad OGC WMS Capabilities response s časovou řadou	39
B.2	Různý způsob vyjádření obsahu elementu Dimension	39
B.3	WMS GetMap požadavek s parametrem TIME	39
C	Příloha C: OGC WMTS – dimenze a časové řady	41
C.1	Příklad OGC WMTS Capabilities response s časovou řadou	41
C.2	WMTS požadavek s parametrem TIME	41
D	Příloha D: OGC WCS 2.0.0 – dimenze a časové řady	43
D.1	Příklad požadavku WCS s časovým omezením - Key Value Pairs	43
D.2	Příklad požadavku WCS s časovým omezením - XML POST	43

E Příloha E: Geometrická validita geoprvků	45
F Příloha F: Kódování češtiny	47
G Příloha G: Formát Atom	49
Literatura	51
Rejstřík	57

Anotace

Tato případová studie shrnuje výsledky konzultací v oblasti otevírání geografických dat, provedených ve prospěch [Institutu plánování a rozvoje hlavního města Prahy \(IPR Praha\)](#), který vznik dokumentu finančně podpořil. Zástupci IPR Praha měli velice kvalifikovaný a ucelený názor na celou problematiku. Konzultace byly zaměřeny na celkový koncept otevírání dat a na technické pasáže, kde byla potřeba většího detailu, doporučení vhodného postupu, případně navržení řešení. Tomu odpovídá i struktura dokumentu, která nemusí být vždy zcela přehledná a dostatečně obecná.

Na základě našich zkušeností z praxe nastiňuje případná technická úskalí a problémy, se kterými je třeba počítat, mají-li být publikovaná data dostupná co nejširšímu okruhu uživatelů. Spíše než jednotlivými nástroji, které jsou v mnoha ohledech věcí osobních preferencí, jsme se zabývali obecnými principy, které je v souvislosti s otevíráním geografických dat potřeba respektovat. V textu jsme se snažili maximálně zúročit naše praktické zkušenosti s různě kvalitními geografickými daty a informačními systémy pro jejich správu. Cílem bylo shromáždit a utřídit přehled témat, které usnadní navržení konkrétního řešení pro otevírání geografických dat.

Licence

Dokument podléhá licenci *Creative Commons 4.0 BY-SA*. Informace o podrobnostech licence najdete na adrese <http://www.creativecommons.cz/>.

Spolupráce

Text studie můžete stáhnout ze serveru GitHub ¹ a dále rozšiřovat standardními postupy (*fork*, *pull request*).

¹ <https://github.com/OpenGeoLabs/otevrena-geodata/>

Revize dokumentu

Datum	Revize	Editoři	Popis
2014-09-30	1.0.0	Jáchym Čepický, Martin Landa, Radek Augustýn, Jan Cibulka, Jan Michálek	První verze dokumentu.
2014-10-03	1.0.1	Radek Augustýn	Drobné revize formátování.
2014-10-05	1.0.2	Jáchym Čepický, Martin Landa, Radek Augustýn, Jan Michálek	Jazyková korektura textu.
2014-11-03	2.0.0	Jáchym Čepický, Martin Landa, Radek Augustýn, Jan Michálek, Lucie Prunarová	Druhá verze dokumentu. Zapracování připomínek IPR Praha.
2014-11-06	2.1.0	Jáchym Čepický	Převod dokumentu do formátu reStructuredTex a jeho publikování na serveru GitHub.
2014-12-01	2.2.0	Jáchym Čepický, Martin Landa, Lucie Prunarová	Revize textu dle připomínek zástupců IPR Praha, seznam zkratk, přeformátování struktury.
2015-01-02	2.2.1	Lucie Prunarová, Jáchym Čepický	Revize textu, doplnění zkratk.
2015-01-21	2.2.2	Lucie Prunarová, Jáchym Čepický	Začlenění připomínek k verzi 2.2.1.

Seznam zkratek

Zkratka	Význam
AOPK	Agentura ochrany přírody a krajiny ČR
ATOM	Atom Syndication Format je webový standard pro publikování syndikovaného obsahu (obsahu, kterému byly zpětně dodány aktuální informace)
DBF	dBase byl první masověji rozšířený systém řízení báze dat
CAD	Z angličtiny computer-aided design, česky počítačem podporované projektování
CSV	Comma-separated values (hodnoty oddělené čárkami) je jednoduchý souborový formát určený pro výměnu tabulkových dat.
CSW	Catalogue Service for Web (přístup k informacím katalogu), OGC specifikace
ČSN	Chráněné označení českých technických norem
ČÚZK	Český úřad zeměměřičský a katastrální
EPSG	European Petroleum Survey Group. Jedná se o široce využívanou databázi zemských elipsoidů, geodetických dat, zeměpisných a kartografických souřadnicových systémů, měrných jednotek, atd.
ETRS89	Evropský terestrický referenční systém, závazný geodetický referenční systém na celém území státu, definovaný technologiemi kosmické geodézie a konstantami, které jsou součástí programů mezinárodních zpracovatelských center, souborem geocentrických souřadnic vybraných bodů geodetických základů, jejichž souřadnice byly vztahy k epoše 1989.
FES	Filter Encoding Specification je standard OGC, který popisuje na systému nezávislou syntaxi pro vyjádření výběru a třídění potřebné při dotazování se na objekty datových sad.
GeoTIFF	Rastrový formát pro geografická data postavený na formátu TIFF
GeoJSON	Vektorový formát pro geografická data využívaný především v prostředí webových aplikací
GIF	Graphics Interchange Format, souborový formát určený pro rastrovou grafiku
GIS	Geografický informační systém
GML	Geography Markup Language je na XML založený formát pro vektorová geografická data, OGC specifikace
GPS	Global Positioning System, česky Globální polohovací systém
HTML	HyperText Markup Language, je značkovací jazyk pro hypertext. Je hlavním z jazyků pro vytváření stránek v systému World Wide Web, který umožňuje publikaci dokumentů na Internetu.
HTTP	Hypertext Transfer Protocol) je internetový protokol určený pro výměnu hypertextových dokumentů ve formátu HTML.

Pokračujte na další stránce

Tabulka 1 – pokračujte na předchozí stránce

Zkratka	Význam
INSPIRE	INfrastructure for SPatial InfoRmation in Europe je iniciativou Evropské komise pro vybudování evropské infrastruktury prostorových informací.
IPR Praha	Institut plánování a rozvoje hl. m. Prahy
ISO	Mezinárodní organizace pro normalizaci
JSON	JavaScript Object Notation (JavaScriptový objektový zápis, JSON) je datový formát určený pro přenos dat
JPEG	Metoda ztrátové komprese používané pro ukládání počítačových obrázků ve fotorealistické kvalitě.
KML	Keyhole Markup Language je aplikací metajazyka XML pro publikaci, distribuci geografických dat
LPIS	Veřejný registr půdy
OGC	Open Geospatial Consortium, mezinárodní standardizační organizace pro oblast GIS a geoprostorových dat
OSM	OpenStreetMap je projekt, jehož cílem je tvorba volně dostupných geografických dat
OWS	Open Web Services, skupina otevřených webových služeb podle standardů OGC
PNG	Portable Network Graphics je grafický formát určený pro bezztrátovou kompresi rastrové grafiky
RÚIAN	Registr územní identifikace, adres a nemovitostí je jedním ze základních registrů ČR
SHP	Formát vektorových dat Esri Shapefile
SDI	Spatial Data Infrastructure
S-JTSK	Jednotná trigonometrická síť katastrální (JTSK, někdy S–JTSK) je pravouhlá souřadnicová síť používaná v geodézii na území České republiky a Slovenska
SOS	Sensor Observation Service, standard OGC OWS (otevřené webové služby podle OGC). SOS umožňuje klientům získat ze serverů různá sensorová data a údaje o pozorování z nich.
SQL	Structured Query Language je standardizovaný dotazovací jazyk používaný pro práci s daty v relačních databázích.
TIFF	Tag Image File Format je jeden ze souborových formátů pro ukládání rastrové počítačové grafiky.
TMS	Tile Map Service, OGC specifikace
TopoJSON	TopoJSON je extenze pro GeoJSON umožňující uložení vektorových dat v topologické formě
URI	Uniform Resource Identifier (jednotný identifikátor zdroje) je textový řetězec s definovanou strukturou, který slouží k přesné specifikaci zdroje informací na Internetu
URL	Uniform Resource Locator (jednotná adresa zdroje) je řetězec znaků s definovanou strukturou, který slouží k přesné specifikaci umístění zdrojů informací na Internetu
VDP	Veřejný dálkový přístup (viz RÚIAN)
VFR	Výměnný formát RÚIAN
VFK	Výměnný formát katastru nemovitostí
Pokračujte na další stránce	

Tabulka 1 – pokračujte na předchozí stránce

Zkratka	Význam
WCS	Web Coverage Service, standard OGC OWS (otevřené webové služby podle OGC), umožňující klientovi požadovat po serveru obrazová (rastrová) data
WFS	Web Feature Service, standard OGC OWS (otevřené webové služby podle OGC), umožňující klientovi požadovat po serveru vektorová data
WGS 84	World Geodetic System 1984, česky Světový geodetický systém 1984
WMS	Web Map Service, standard OGC OWS (otevřené webové služby podle OGC), umožňující klientovi požadovat po serveru vyrenderované obrázky mapových kompozic
WMTS	Web Map Tile Service, OGC specifikace
WWW	World Wide Web, soustava propojených hypertextových (na sebe navzájem odkazujících) dokumentů
XML	Extensible Markup Language (rozšiřitelný značkovací jazyk) je obecný značkovací jazyk, který byl vyvinut a standardizován konsorciem W3C
XML-RPC	Protokol, s jehož pomocí lze velice jednoduše provádět vzdálené volání procedur.
XSD	XML Schema XSD je jedno z XML schémat, jazyků pro popis XML.
XSLT	eXtensible Stylesheet Language Transformations (Transformace XSLT) slouží k převodům zdrojových dat ve formátu XML do libovolného jiného požadovaného formátu
ZCHÚ	Zvláště chráněná území

Úvod

Je všeobecně uznávaným faktem, že sebelepší řešení není udržitelné, pokud není naplněno kvalitními a aktuálními daty. V poslední době dochází v této souvislosti v čím dál větší míře ke sdílení a šíření dat prostřednictvím sítě Internet. Pojmy jako *“linked data”* nebo *“sémantický web”* nejsou dávno jen teoretickým konceptem, ale aktuální skutečností. Bez nadsázky lze říci, že data jsou jedna z komodit, které podporují jak technický pokrok, tak rozvoj občanské společnosti. Tento fakt začíná být po vzoru našich nejbližších sousedů, především Německa, ale i Rakouska či Velké Británie reflektován i veřejnou správou v České republice v podobě postupného otevírání dat.

Právě ve Velké Británii vznikly dvě celosvětově nejvýznamnější hnutí v otevírání dat. Prvním z nich je nezisková organizace [Open Knowledge Foundation](#) (OKFN), která se orientuje na propagaci otevřenosti ve znalostech a datech obecně. OKFN stojí mimo jiné za projektem katalogu CKAN [ref45] a tzv. Open Data Index (žebříček otevřenosti dat). Tento žebříček shrnuje úroveň států z pohledu otevřenosti a dostupnosti nejdůležitějších datových sad jako jsou např. jízdní řády, státní rozpočet, výsledky voleb, obchodní rejstřík a podobně. Základní zkušenosti s právními, sociálními a technickými aspekty OKFN shrnuje publikace *Open Data Handbook* [ref70], ze které částečně čerpá i tato studie. Druhým z uvedených hybatelů je projekt [OpenStreetMap](#) (OSM), který byl založen v roce 2006 právě ve Velké Británii s cílem komunitního vytváření volně dostupných geografických dat a následně jejich vizualizace do podoby silniční mapy, uličního plánu měst a dalších výstupů. Vzhledem k tomu, že je od počátku založen na kolektivní spolupráci a na koncepci Open Source, se rychle rozšířil do celého světa. Data z projektu OSM jsou poskytována pod otevřenou licencí Open Database Licence.

Veřejná správa vždy shromažďovala data potřebná pro výkon svých agend. Zpřístupnění těchto dat umožní veřejnosti nejen využít informace v nich obsažené, ale především *řádově znásobit jejich hodnotu* tím, že kdokoliv nad nimi bude moci bez jakéhokoliv omezení postavit vlastní aplikaci či službu, anebo je podrobit nezávislé analýze. Právě proto je důležité, aby co nejvíce dat vytvářených veřejnou správou bylo publikováno a volně poskytováno široké veřejnosti k dalšímu užití. Jen tak může být zcela beze zbytku využít potenciál dat sbíraných z veřejných prostředků.

Otevírání dat je proces, který se týká několika oblastí. Kromě technických záležitostí se jedná především o licenční politiku, která v ideálním případě umožní vývojářům navazujících aplikací data předávat dál a vytvářet nad nimi další odvozená díla, která nebudou zatěžována žádnými restrikcemi.

Obsahem této studie jsou zejména geografická data (tzv. geodata), která se ve velké míře řídí na rozdíl od ostatních typů dat vlastními zvyklostmi, standardy či technickými normami. Proces standardizace je v oblasti geoinformačních technologií, na rozdíl od jiných technologických oblastí, podporován již delší dobu, což výrazně ulehčuje tvorbu dalších doporučení. V současnosti drtivá většina výrobců softwarových produktů (proprietárních i otevřených), ale např. i Evropská komise skrze směrnici INSPIRE navazuje na existující standardy v této oblasti, které pomáhá dále rozvíjet. Standardy jsou vytvářeny mezinárodní standardizační organizací [Open Geospatial Consortium](#) (OGC) a technickou komisí [ISO TC 211](#). Mezinárodní technické normy ISO mohou následně být přejímány jako normy národní (označené

jako ČSN EN ISO, například ČSN EN ISO 19136 *Geografická informace – Značkovací jazyk geografie (GML)* [ref77]).

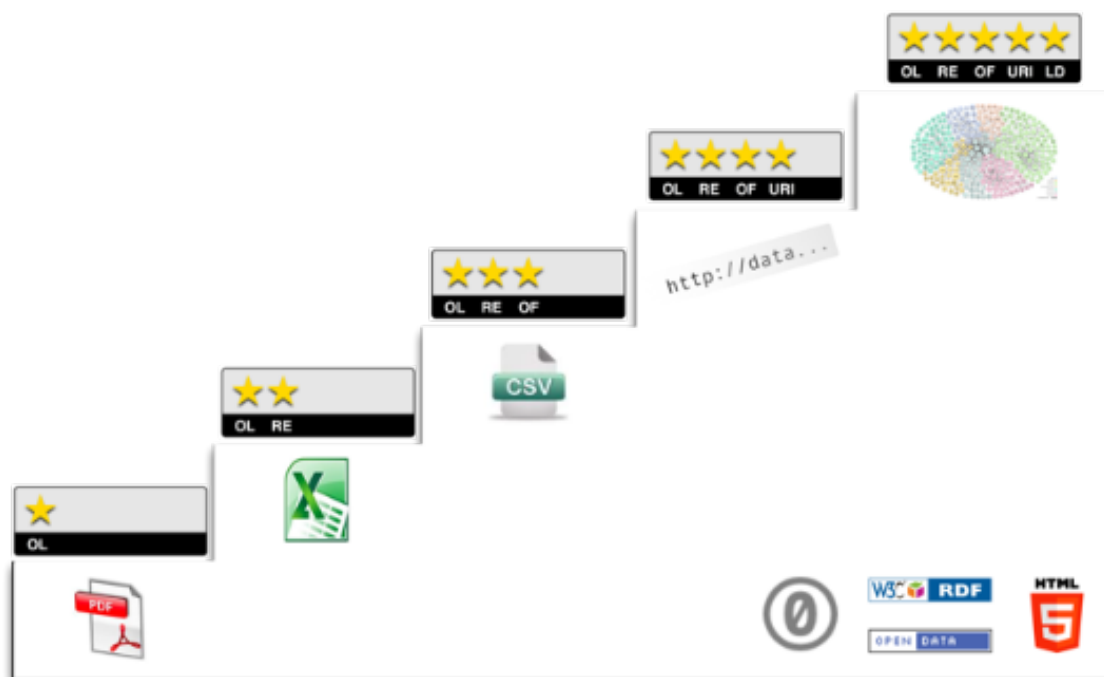
1.1 Otevřená data

Aby data mohla být označena za *otevřená*, musí splňovat několik legislativních a technických podmínek (viz níže). Tyto podmínky nejsou v čase konstantní, vyvíjí se spolu s technickými prostředky, standardy ale i legislativou či postoji společnosti.

V současnosti těmito podmínkami myslíme zejména:

- Data jsou dostupná na síti Internet a dohledatelná běžnými nástroji
- Data jsou strojově čitelná, tedy v textovém či binárním formátu
- Data musí být úplná, tj. jsou zveřejněná v takovém rozsahu, aby nechyběly některé fragmenty umožňující jejich využití. Data by zároveň měla být zveřejněna v maximálním možném rozsahu. //musí být úplná vs. zveřejněna v maximálním možném rozsahu - nejasný význam vět//
- Legislativní a technické překážky pro využití dat jsou minimální

Tim Berners-Lee [ref15] sestavil tzv. *pětihvězdičkový systém* hodnocení otevřenosti dat, jehož smyslem je umožnit jednoduchou orientaci ve stupni otevřenosti datových sad. Čím výše se datová sada dostane, tím větší je možnost jejího využití veřejností.



Obrázek 1.1: Pětihvězdičkový systém otevřených dat podle Tim Berners-Lee [ref15]

★ Dostupná data

Data jsou zveřejněna na Internetu, nezáleží na formátu. Data, která jsou zveřejněna na síti Internet v jakémkoli formátu, ale pod otevřenou licenci pro jejich využití. Existují tedy určitá technická omezení pro jejich využití, nicméně legislativní omezení jsou odstraněna. Příkladem může být mapové dílo ve formátu PDF.

★★ Strukturovaná data

Data jsou zveřejněna ve strojově čitelném formátu. Data jsou zveřejněna ve strojově zpracovatelném formátu, který není otevřený. Příkladem mohou být data ve formátu Microsoft Excel nebo data uložená v Esri GeoDatabase.

★★★ Data v otevřeném formátu

Data jsou zveřejněna v otevřeném formátu. Strojový formát, ve kterém jsou data zveřejněna, je otevřený, tj. veřejně publikovaný s licenci umožňující jeho využití. Příkladem jsou například formáty Geospatial Markup Language (GML), OGC GeoPackage nebo Esri Shapefile

★★★★ Data s identifikátorem

Data jsou dohledatelná, mají unikátní identifikátor. Data jsou opatřena identifikátorem Uniform Resource Identifier (URI), unikátním v rámci celé sítě Internet. Jednotlivé prvky datové sady by měly být identifikované pomocí URI ve tvaru použitelném pro protokol HTTP tzn. v podobě adresy URL. Tím se zajistí, že uživatelé mohou daný prvek kdykoli najít.

★★★★★ Provázaná data

Data jsou navzájem prolinkovaná, lze mezi nimi navigovat, "surfovat"[\[ref71\]](#). Data jsou nejen identifikovatelná pomocí URI, ale obsahují odkazy na další datové sady. Stejně jako datové sady jsou pomocí odkazů provázány i jednotlivé prvky z datových sad. Data jsou dále opatřena popisnými informacemi (metadaty) tak, aby v nich bylo možno jednoduše vyhledávat.

1.1.1 Výhody a omezení otevírání dat

Je potřeba zdůraznit, že proces otevírání dat je oproti stávajícímu stavu navíc vynaložená práce. Ta je požadována zejména po poskytovateli dat, který, chce-li se na určitou úroveň otevřenosti dat dostat, musí do stávajících datových sad dodat požadované chybějící informace.

V této části textu rozebereme pozitiva otevírání dat a požadavky kladené na poskytovatele otevřených dat, i jejich uživatele.

Výhody ✓ a omezení ✗ – z hlediska uživatele

★

- ✓ Data lze prohlížet
- ✓ Data lze tisknout
- ✓ Data lze uložit na lokální disk
- ✓ Data lze vložit do dalšího systému nebo databáze
- ✓ Data lze podle potřeby a možností použitého formátu měnit, doplňovat či odvozovat další produkty
- ✓ Data je možné sdílet s dalšími uživateli

★★

- ✓ Data lze automaticky zpracovávat pomocí proprietárního software
- ✓ Data lze vyexportovat do jiného formátu

★★★

- ✓ Uživatel může manipulovat s daty, aniž by byl závislý na vlastnictví konkrétního, většinou proprietárního software

★★★★

- ✓ Uživatel může na data odkazovat z kteréhokoli jiného místa na webu
- ✓ Lze pořizovat trvalé odkazy na data
- ✓ Lze znovu použít část dat, aniž by se musela vytvářet jejich fyzická kopie
- ✓ Lze kombinovat datové sady mezi sebou, protože použité URI jsou vždy unikátní
- ✗ Porozumění struktuře takto publikovaných otevřených dat je komplikovanější, než pochopení jednoduché tabelární nebo stromové struktury

★★★★★

- ✓ Lze najít další datové sady při procházení té současné
 - ✓ Datové sady jsou publikovány v jasně definovaném schématu
 - ✗ Možnost výskytu neexistujících cílů, které je potřeba ošetřit, podobně jako když na webovém serveru neexistuje požadovaná adresa
 - ✗ Prezentace dat z externích zdrojů jako ověřený fakt je riskantní
-

Výhody ✓ a omezení ✗ – z hlediska poskytovatele

★

- ✓ Odpadá opakovaná činnost související s distribucí dat, uživatelé si je mohou stáhnout z Internetu
- ✓ Vlastní otevření je většinou snadné a s minimálními náklady

★★

- ✓ Otevření dat i uživatelům, kteří mají zájem kromě jejich prohlížení také o jejich zpracování. Tím se výrazně zvětšuje velikost skupiny uživatelů, kteří budou takto publikovaná data využívat
- ✓ Vlastní otevření je většinou snadné a s minimálními náklady

★★★

- ✓ Další rozšíření skupiny potenciálních uživatelů o ty, kteří nepoužívají programové vybavení kompatibilní s vybavením poskytovatele
- ✗ Náklady na konverzi a uložení dat do zvolených formátů. Tyto náklady nemusí být v konečném důsledku příliš vysoké, protože konverze mezi jednotlivými formáty je poměrně běžná a dobře zvládnutá. Je však třeba s nimi počítat.

★★★★

- ✓ Uživatel má velice dobrou kontrolu členění dat a může optimalizovat přístup k nim
- ✓ Jiní poskytovatelé se mohou na data navázat a tím je zlepšit na úroveň ★★★★★
- ✗ Většinou je nutný zásah do struktury dat
- ✗ Poskytovatel musí přiřadit URI k datům a zabezpečit jejich prezentaci
- ✗ Poskytovatel musí najít existující postupy nebo vytvořit své vlastní

★★★★★

- ✓ Data jsou dohledatelná a prohledatelná, čímž se výrazně zvyšuje jejich hodnota
- ✓ Poskytovatel může profitovat ze vzájemného provázání dat, podobně jako uživatelé
- ✗ Poskytovatel musí investovat do propojení vlastních dat s dalšími datovými sadami
- ✗ Poskytovatel musí udržovat tato propojení aktuální, pokud možno odstraňovat propojení vedoucí na již neexistující cíle (např. pokud je nějaká webová stránka odstraněna).

1.1.2 Pozitivní aspekty otevírání geodat, best practices

Otevřením dat, tj. jejich zpřístupněním široké veřejnosti bez dalších omezení, dochází mimo jiné ke zvyšování jejich hodnoty. Většina autorů se shoduje na tom, že státní a veřejné instituce již pouhým vystavením dat na Internetu jejich původní hodnotu násobí (např. [ref01] str. 19, [ref78]).

Dále se ukazuje, že s nejzajímavějšími případy využití dat přijdou spíše často komerční společnosti nebo jednotlivci než instituce, které data původně vytvořily anebo zadaly jejich vytvoření. Tento efekt byl znatelný například na akci “První Pražský GeoData Hackathon” [ref02]. Během jednoho víkendu vzniklo nad vybranými datovými sadami publikovanými Institutem plánování a rozvoje hlavního města Prahy šest zcela nových aplikací a analýz [ref03] a to i přes počáteční technické problémy související s formou zpřístupnění geodat pomocí proprietárního rozhraní XML-RPC. Jako příklad ze zahraničí lze uvést například město Londýn, které otevřelo data spojená s přepravním systémem. Na základě těchto dat následně vyrostla řada aplikací pro plánování cesty či domlouvání spolujízd a další.

Otevírání dat má v obecné rovině hned několik pozitivních efektů, a to zejména:

- Kvalitnější služby pro občany
- Zvýšení informovanosti obyvatelstva
- Podpora občanské společnosti
- Zvýšení ekonomického potenciálu
- Zlepšení transparentnosti veřejné správy
- Efektivnější veřejná správa
- Redukce administrativy a zátěže IT infrastruktury veřejné správy

Kvalitnější služby, zvýšení informovanosti veřejnosti

Projekty jako “Mapa kriminality” [ref04], různé cenové srovnávače (např. server <http://www.heureka.cz>), “Váš majetek” [ref05] a další vedou k efektivnějšímu nakládání se soukromým a především veřejným majetkem. Soukromé osoby, ale i veřejné instituce dostávají k dispozici větší množství relevantních podkladů pro efektivnější rozhodování. Dochází tak i k propojení veřejných institucí s občanskými iniciativami, které pak zpětně pozitivně ovlivňují veřejnou správu. Jako příklad můžeme uvést Policii ČR, která využívá vizualizační možnosti aplikace “Mapa kriminality”.

Obdobným příkladem je Datablog Hospodářských novin, který použil data z Pražského dopravního podniku a na základě nich vytvořil aplikaci zobrazující dostupnost jednotlivých městských částí prostřednictvím Pražské hromadné dopravy [ref10]. Tato aplikace umožňuje například srovnat dostupnost městské hromadné dopravy v Praze před a po velké změně jízdního řádu k 1.9.2012.

Projekt “What’s the Plan” [ref13], který byl vytvořen v rámci projektu “EU Plan4Business”, kombinuje různé otevřené datové zdroje mimo jiné data z Registru Územní Identifikace, Adres a Nemovitostí (RÚIAN) [ref12] a z databází Ministerstva financí [ref07]. Jeho účelem je zobrazení přehledných informací vhodných zejména pro potenciální investory, kteří hledají pro svůj investiční záměr v Evropě vhodnou lokalitu.

Zvýšení transparentnosti veřejné správy

Přehledy hospodaření obcí v ČR jsou dostupné na webových stránkách “Rozpočet obce” [ref06]. Tato aplikace vznikla na základě dat, která v surové podobě publikuje Ministerstvo financí prostřednictvím

registru ÚFIS [ref07]. Na základě analýz těchto dat již vznikají další akademické studie [ref08], [ref09].

Některé politické strany zveřejňují své finanční zprávy na Internetu. Pomocí automatických nástrojů (např. [ref11]) lze zpracovat a vizualizovat mimo jiné i jejich aktuální finanční situaci.

Jako další příklad lze uvést zveřejňování územních plánů, díky kterému dochází k většímu zapojení veřejnosti do diskuze o budoucnosti obce [ref14].

Redukce administrativy a zátěže IT infrastruktury veřejné správy

Veřejná správa je často zatěžována požadavky veřejnosti na poskytnutí dat, která vytváří a spravuje. Pro vyřízení těchto žádostí musí udržovat kapacity a IT infrastrukturu. Právní oddělení musí být připraveno na posouzení žádostí ze strany veřejnosti a případně zdůvodnit jejich neposkytnutí atp. Po veřejné správě přitom ale nejsou vyžadovány žádné dodatečné analýzy, ale pouze kopie již existujících datových sad. Každá žádost je posuzována a zpracovávána individuálně, většinou specializovanými pracovníky. Také to zvyšuje celkovou administrativní zátěž.

Vzhledem ke složitosti tohoto způsobu získávání dat je většinou uživateli vyžadováno více dat než skutečně potřebují, zvláště pokud jsou poskytována bezplatně. Zveřejněním vybraných datových sad na webových stránkách příslušné složky veřejné správy tato zátěž odpadá nebo se výrazně minimalizuje. Dalším faktorem, který snižuje zátěž IT infrastruktury veřejné správy je skutečnost, že uživatelé mohou přistupovat pouze k datům, která doopravdy potřebují. To se projeví jak při stahování datových sad, tak v případech prohlížení rastrových i vektorových služeb.

1.1.3 Klíčové příklady otevřených geodat

Úspěšných příkladů, ze kterých je možné čerpat, je velké množství. Zřejmě nejkompletnější a nejlépe udržovaný přehled otevřených geografických datových sad v České republice je udržován na Wiki stránce české části projektu OpenStreetMap [ref52].

V této kapitole si představíme některé příklady otevírání geografických dat u nás i v zahraničí (u příkladu uveden stupeň otevřenosti dat dle pětihvězdičkového systému):

- Londýn
- Chicago
- Berlín
- Vídeň
- Děčín
- Opava
- Český statistický úřad
- Český úřad zeměměřický a katastrální
- Ministerstvo zemědělství

Londýn ★★★

Vedení města Londýn si dalo za cíl zveřejňovat co nejvíce dat, které sbírá a uchovává z titulu své funkce. Výsledkem je London Datastore [ref75], který slouží k poskytování dat veřejnosti, výhradně

zdarma. Na svých stránkách vybízí k používání těchto dat a k vytváření aplikací nad nimi. Metadata k datovým sadám jsou dostupná v jednoduchém formátu Comma Separated Values (CSV). Na základě těchto dat následně vznikla řada aplikací, například pro plánování cesty či domlouvání spouhů. Přehled aplikací vzniklých v Londýně, a to nejen nad dopravními daty, je k dispozici na adrese <http://data.london.gov.uk/datastore/inspirational-uses>.

Chicago ★★★

Město Chicago publikovalo některá svá geodata pomocí portálu GitHub [ref57] a umožnilo tak uživatelům pohodlný přístup k jejich stažení. Vedle samotných dat město publikovalo i analytické skripty v jazyce R s tím, že je zájemci z řad veřejnosti mohou otestovat a případně navrhnout jejich zlepšení či opravit chyby.

Berlín ★★★

Město Berlín publikuje na adrese <http://daten.berlin.de> volně dostupné datové sady vesměs pod licencí Creative Commons [ref47]. Geodata jsou publikovaná pomocí proprietárního REST rozhraní, kde pro přenos dat využívá formáty XML a JSON. Město Berlín dále nabízí WFS službu a Atom na webovém portálu územního plánování <http://www.stadtentwicklung.berlin.de/geoinformation/geodateninfrastruktur/index.shtml>.

Vídeň ★★★

Město Vídeň publikuje svá otevřená data na adrese <https://open.wien.at>. Geodata jsou poskytována pomocí webových služeb OGC. Pro tabelární data je používán jednoduchý formát CSV.

Děčín ★★★

Město Děčín ve spolupráci s Fondem Otakara Motejla a Fórem pro otevřená data jako první v České republice otevřelo svá data. Data jsou zveřejněna na webových stránkách statutárního města Děčín na adrese http://www.mmdecin.cz/dokumenty/cat_view/238-otevrena-data a označena jako "otevřená data". Uživatelé je povoleno data dále šířit a komerčně využívat za podmínky uvedení zdroje. Data publikovaná ve skupině geografických a mapových dat jsou dostupná ke stažení ve formátu SHP. Licence těchto dat není nikde na stránkách města zmíněna, lze tedy soudit, že se jedná o tzv. *public domain* – veřejné dílo, které je volně k použití bez specifické licence.

Opava ★

Podobnou cestou jako Děčín se vydalo město Opava, které na svých stránkách na adrese <http://www.opava-city.cz/cs/mapovy-portal> publikuje data ve formě prohlížečích služeb WMS [ref20] a některé mapy ve formátu PDF.

Český statistický úřad ★★★

Český statistický úřad zveřejňuje na základě usnesení vlády dostupná data z výsledků voleb v otevřených formátech, zároveň s textovým popisem obsahu zveřejněných dat na adrese http://www.czso.cz/csu/redakce.nsf/i/otevrena_data. Dále zveřejňuje data ze Sčítání lidu domů a bytů 2011, která jsou k dispozici ve formátech CSV, DBF a XML.

Český úřad zeměměřický a katastrální ★★★

K příkladům “*best practice*” přístupů pro otevírání geodat zcela jistě patří aplikace Veřejného dálkového přístupu (VDP) k datům RÚIAN (Registr Územní Identifikace, Adres a Nemovitostí) provozovaná Českým úřadem zeměměřickým a katastrálním (ČÚZK) na adrese <http://vdp.cuzk.cz/>. Jedná se o projekt v českém kontextu jedinečný jak svým rozsahem, tak především objemem a kvalitou vydávaných geodat a mírou jeho využívání veřejností.

Na tomto příkladu lze demonstrovat zásadní problémy, které je potřeba pro úspěšné otevření geodat vyřešit. Začít můžeme obligátním “kvalitní řešení nemůže existovat bez kvalitních dat”. V tomto případě jsou data poskytovaná prostřednictvím VDP jako předgenerované soubory ve Výměnném formátu RÚIAN (VFR). Data jsou rozčleněna podle územních jednotek a dále poskytovaná v různých verzích podle typu a úrovně detailu poskytované informace. Zásadním aspektem je i to, že jsou poskytována ve dvou formách, jako celá datová sada aktuální k prvnímu dni v měsíci a dále jako denní změnové soubory. Tento přístup publikování je velmi výhodný, dochází k snížení zátěže IT infrastruktury na straně poskytovatele geodat a současně k zlepšení komfortu cílových uživatelů, kteří si mohou jednoduše udržovat svoji kopii datové sady v aktuálním stavu pravidelným začleňováním změnových souborů. Díky tomu vznikl kolem těchto dat velmi rychle relativně bohatý ekosystém nástrojů a služeb. K úspěšnosti aplikace přispěla také kvalitní dokumentace publikovaná s předstihem před spuštěním VDP.

Z technického hlediska je formát VFR odvozen ze standardu OGC GML [ref19], proto je na straně aplikací poměrně snadno uchopitelný. To umožnilo rychlé začlenění podpory tohoto formátu do nejrozšířenějších Open Source a proprietárních GIS nástrojů. Zároveň umožňuje vyjádřit vazby mezi jednotlivými prvky s ohledem na strukturu datového modelu, ze kterého RÚIAN vychází. Je potřeba zmínit, že výměnný formát VFR není standardizovaný mezinárodní institucí, ale ze standardu (GML) vychází, je dobře dokumentovaný a jeho licence je nerestriktivní, proto se jedná o formát otevřený, i když proprietární.

Ministerstvo zemědělství ★★★

Projekt veřejný registr půd LPIS [ref73] provozovaný Ministerstvem zemědělství představuje novou filozofii v řešení GIS pro evidenci využití zemědělské půdy. Jeho základním cílem je umožnit bezproblémové zvládnutí administrace a kontroly žádostí o zemědělské dotace. LPIS poskytuje i mapové služby, které lze vyhledat v metadatovém katalogu. Mezi jinými je k dispozici i webová služba WFS [ref21].

1.2 Uživatelé otevřených geodat

Uživatele lze obecně dělit podle různých kritérií: na základě schopností, zkušeností, specializace, požadavků na data a dalších. Tato kritéria jsou vzájemně provázána, proto se nevyhneme zjednodušení na modelové skupiny uživatelů.

Pro otevírání geografických dat se jako klíčové jeví tyto skupiny uživatelů:

- Uživatel mapového portálu
- Běžný uživatel geografického informačního systému (GIS)
- Specialista GIS (geoinformatik)
- Datový analytik (mimo obor GIS)
- Programátor/vývojář

1.2.1 Uživatel mapového portálu

Největší částí uživatelů geografických dat je laická veřejnost, která využívá některý z mapových portálů, případně specializované nástroje pro práci s mapou. Tato skupina uživatelů není přímo cílovou skupinou otevírání geografických dat, informace se k nim povětšinou dostanou zprostředkovaně právě pomocí portálu, který využívají. Potřeby a očekávání těchto uživatelů bereme v úvahu zejména při propagaci otevřených dat.

1.2.2 Běžný uživatel geografického informačního systému (GIS)

Tento typ uživatele obvykle zvládá řešení základních, až středně obtížných úloh s využitím desktopového GIS, případně nástroje na tvorbu výkresů CAD (AutoCAD, Microstation apod). Je schopen pracovat s daty, která lze snadno připojit nebo jednoduše importovat v konkrétním programu, se kterým pracuje. S komplexnějšími formáty a datovými sadami si obvykle poradí pouze tehdy, pokud se dají otevřít nástrojem integrovaným do tohoto programu. Data, která sám nevytváří používá obvykle jako podkladová, případně pro provádění analýz, často jednorázově. Jeho zájmem jsou tedy spíše data platná k určitému datu, než data průběžně udržovaná v aktuálním stavu z webové služby.

Příkladem tohoto typu uživatele jsou studenti ne-geoinformatických oborů, vědci, tvůrci studií, odborníci vytvářející tematická data, plány či mapy (např. lesní plány), stavebníci, architekti, pracovníci ochrany přírody a další odborníci, kteří využívají GIS jako jeden z více pracovních nástrojů.

Obecně lze tuto skupinu uživatelů považovat za poměrně konzervativní, GIS není jejich primárním zaměřením, obvykle se jednorázově naučí základům problematiky a své znalosti dále příliš nerozvíjí. Dalším omezením, zejména pokud využívají proprietární software, může být využívání starých verzí, které ještě nemají podporu novějších formátů.

Pro rastrová data je pro tuto skupinu uživatelů ideální využívat služby WMS, případně WMTS. U vektorových dat je nejvhodnější stahování jednotlivých vrstev v jednom z rozšířených souborových formátů (např. SHP). Použití WFS je ideální cílový stav, ale může být problematické při větších objemech dat, který může uživatel (i když ne zcela vědomě) po serveru požadovat. Řešením je "stránkování" odpovědi služby, tato možnost ale bohužel není na všech klientských programech běžně implementována.

1.2.3 Specialista GIS (geoinformatik)

Specialista GIS je odborníkem pro práci s geografickými daty, který preferuje ucelenou datovou sadu včetně popisných informací, zajímá se o aktuálnost a přesnost dat. Nemá problém vybrat z nabídky tu, která nejvíce odpovídá jeho potřebám. Kromě pokročilých analýz a vizualizací připravuje datové sady pro použití v rámci organizace jak pro vnitřní informační systémy, tak pro méně zkušené uživatele.

Překážkou, která se u této skupiny objevuje, je přílišná vazba na konkrétní software, se kterým pracují. Specialisté se vyhýbají použití nového, pro ně neznámého software a nebo jsou v situaci, že jim zaměstnavatel či pracovní podmínky či oborové zvyklosti neumožní alternativní software využívat.

Specialisté si obvykle vytváří vlastní databáze, ve kterých se snaží udržovat aktuální kopii celých datových sad, což umožňuje provádět rychlé a komplexní analýzy nad velkými daty. Poskytují webové služby s vysokou dostupností dimenzované pro potřeby konkrétních aplikací. Pro udržování aktualizované datové sady je nejvýhodnějším způsobem distribuce poskytování stavových dat a změnových souborů, ideálně opatřených kvalitními metadaty ve standardním formátu. V případě poskytování dat touto cestou je potřeba, aby bylo umožněno automatizované vytěžování dat, tedy umístění souborů na odvoditelné adrese nebo předávání odkazů na soubory ke stažení.

1.2.4 Datový analytik

Další skupinou uživatelů je odborník na analýzu dat bez znalosti GIS. Jelikož nezná běžně se vyskytující datové formáty ani pravděpodobně nemá k dispozici potřebný software, bývá odkázán na vlastní nástroje (nejčastěji databáze a statistický software), pomocí nichž data zkoumá.

Datový analytik tedy potřebuje formát otevřený, ideálně textový, dobře dokumentovaný. Práci usnadní zejména formalizované popisy dat, jako jsou například XSD schémata, která se používají při generování různých šablon pro dokumenty XML, připojení dat apod.

Ideálním způsobem distribuce jsou pro něj předgenerované soubory v textovém strukturovaném formátu (JSON, XML, ...). Je ale schopen, pokud je dostupná uživatelská dokumentace, využít i webové služby.

1.2.5 Programátor/vývojář

Přestože skupina vývojářů není příliš početná, je velice důležitá. Vývojáři totiž zpřístupňují data pomocí aplikací, obvykle vyvinutých na míru specifickým cílům nebo tematickému okruhu uživatelů. Kvalita popisu formátu a dostupnost knihoven pro jejich využití značně ovlivňuje náklady, se kterými tyto aplikace vznikají. Přestože sami jsou vývojáři často skalními zastánci různých technologií (.NET, C++, Java atd), geografická data jsou pro ně cizí a akceptují proto jakýkoli funkční a dobře popsany standard, nejlépe přímo s knihovnou k jeho využívání. Důležitým faktorem pro tuto skupinu je také otevřená licenční politika, tedy možnost poskytnutá data přizpůsobovat konkrétním potřebám. Dále oceňují stabilitu poskytovaných služeb a dat, protože připravují služby určené k dlouhodobému využití. Je proto vhodné, aby data byla poskytována dlouhodobě a bez zásadních změn.

Vývojáři sledují trendy a vývoj v oboru Informačních technologií (IT), mají přehled o novinkách a jsou schopni využít výhod nových a zatím nepříliš rozšířených forem přenosu dat i technologií. Ani komplikovanější formáty pro ně nemusí být problémem, pokud je dostupná kvalitní dokumentace.

Ideálním způsobem distribuce dat pro takto vymezenou skupinu uživatelů jsou proto bezesporu webové služby, pokud jsou ovšem stabilní, rychlé a vhodně nastavené. S jejich využitím velice efektivně vytvoří aplikaci, aniž by musel provozovat vlastní server a udržovat na něm aktuální data.

Technické a právní aspekty otevírání datových sad

2.1 Výběr licence pro otevřená geodata

Poznámka: Autoři následující kapitoly nejsou právníci, pouze popisují své zkušenosti v dané oblasti. Před aplikací doporučení z tohoto textu je proto důrazně doporučeno jej konzultovat s expertem na autorské právo.

Otevíraná data by měla být licencovaná tak, aby na straně jedné zbytečně neomezovala uživatele a na straně druhé dostatečně chránila práva poskytovatele dat. Podstatné podmínky tedy jsou:

- Uživatel smí data kopírovat a využívat v rámci své činnosti, rovněž smí poskytnutá data začlenit do svého díla
- Uživatel nesmí přeprodávat původní data
- Uživatel musí vždy uvést zdroj dat a licenci; nesmí tak ale činit způsobem, který naznačuje, že je podporován poskytovatelem dat

Od nového roku upravuje problematiku licenčních smluv nový občanský zákoník (část čtvrtá, hlava druhá, díl druhý, oddíl 5). Podstatný je § 2373 odst. 1), který umožňuje projevit vůli k uzavření smlouvy i směrem k *“neurčitému počtu osob”*. Dále tento odstavec obsahuje ustanovení, že *“Obsah smlouvy nebo jeho část lze určit také odkazem na licenční podmínky, jež jsou stranám známé nebo veřejně dostupné.”*

Následující paragraf pak umožňuje vyjádřit *“souhlas s návrhem na uzavření smlouvy provedením určitého úkonu bez vyrozumění navrhovatele, zejména poskytnutím nebo přijetím plnění.”*

Poskytovatel může data opatřit licenčními podmínkami, které uživatel přijme samotným užíváním díla. Podmínkou je, aby byl text licence uživateli dat volně dostupný.

Tuzemská státní správa už tento způsob licencování místy využívá, jde konkrétně o licenci Creative Commons [ref47]. Obsah pod touto licencí zveřejňují Český hydrometeorologický úřad, Ústav zdravotnických informací a statistiky ČR či Celní správa.

2.2 Otevřené formáty, webové služby a distribuce otevřených geodat

Tato kapitola se zabývá formáty otevřených geodat. Jsou zde popsány populární datové formáty, které jsou ale pro účel otevírání dat nevhodné i nové, rychle se rozvíjející vhodné formáty. Podrobněji jsou zde zhodnoceny dostupné webové služby jako vhodný a široce využívaný způsob distribuce geodat.

Formát určuje význam dat v elektronickém souboru. Formát je způsob, jakým jsou data v souboru uložena. Formáty jsou často navrženy specificky pro určitý typ dat. V základu lze formáty rozdělit na binární (z pohledu člověka se jedná o souvislý tok hodnot 0 a 1) a textové (které jsou pro člověka poměrně jednoduše čitelné). Příkladem binárního formátu může být JPEG speciálně navržený k ukládání obrazových dat. Příkladem textového formátu může být formát JSON či XML určený k popisu a ukládání libovolných stromových datových struktur.

Webové služby jsou softwarové systémy umožňující interakci dvou strojů v počítačové síti. Počítače mezi sebou komunikují pomocí strojově zpracovatelného formátu zpráv. Webové služby ke své komunikaci využívají protokolu HTTP (Hypertext Transfer Protocol), o který je opřena World Wide Web (WWW – odtud webové služby). Jeden z počítačů ve vzájemné komunikaci vždy vystupuje jako klient (požaduje po serveru splnění nějakého úkolu), druhý jako server (vyřizuje žádost a posílá klientovi odpověď). Výsledkem odpovědi může například být i obrázek v patřičném souborovém formátu, jako je např. JPEG.

2.2.1 Legislativní pravidla otevírání dat - INSPIRE

Kromě souvisejících zákonů a nařízení je pro otevírání geodat relevantní zejména evropská směrnice INSPIRE [ref26] ze dne 25. dubna 2007, která vstoupila v platnost 15. května 2007. Tato směrnice tvoří evropský legislativní rámec potřebný k vybudování evropské infrastruktury prostorových informací zejména k podpoře environmentálních politik a politik, které životní prostředí ovlivňují. Hlavním cílem směrnice INSPIRE je poskytnout množství kvalitních a standardizovaných prostorových informací. Směrnice byla transponována do národní legislativy České republiky novelou zákona č. 380/2009 Sb. [ref27].

Otevřené formáty

Formáty pro publikování otevřených geodat by měly být zejména otevřené – ať už proprietární, tj. spravované konkrétní firmou anebo standardizované konsorciem OGC či technickou normou ISO. Otevřený formát je takový formát, ke kterému existuje volně dostupná dokumentace a zároveň jeho licence umožňuje jeho volné využívání. V ideálním případě je vyvíjen nezávislou mezinárodní standardizační organizací a není navázán přímo na žádný privátní subjekt.

Z dlouhodobého hlediska je výhodnější využívat formáty otevřené a standardizované, než proprietární (jakkoliv mohou být široce rozšířené mezi uživateli a jimi používanými programy), zejména protože:

- Stává se, že firma (autor proprietárního formátu a držitel licenčních práv k němu) svůj formát změní, případně programové nástroje třetích stran pak reagují se zpožděním, pokud vůbec
- Uživatelé jsou nuceni používat konkrétní software, což nemusí být pro některé z nich z technických, licenčních nebo finančních důvodů akceptovatelné
- Celý systém se stává závislý na jedné firmě či dodavateli softwarového řešení, je nepřenositelný a v budoucnu jen těžko rozšiřitelný

Z pohledu výše zmíněné *pětihvězdičkové konvence*, můžeme pro geodata uvést následující příklady vhodných datových formátů:

★	Tisknutelná mapa je uložena ve formátu PDF nebo v rastrovém formátu jako obrázek, například JPEG, GIF či PNG
★★	Vektorová data jsou uložena v uzavřeném proprietárním formátu, například DWG firmy Autodesk
★★★	Vektorová data jsou uložena ve formátu SHPP který je sice proprietární, ale je otevřený; nebo jako GML, což je otevřený standard a zároveň technická norma ISO
★★★★	Data jsou opatřena metadaty a jednotlivé datové sady jsou jimi navzájem provázány. Předpokládá se existence metadatového katalogu, založeného na standardu OGC CSW, ve kterém lze data a služby vyhledávat
★★★★★	Jednotlivé soubory datové sady avzájem provázány pomocí unikátních identifikátorů

2.2.2 Formáty souborů pro distribuci otevřených geodat

Pro ukládání, zpracování a výměnu geodat existuje velké množství formátů. Z hlediska otevřených dat se prozatím můžeme omezit na formáty rastrové a vektorové.

Formát GeoTIFF [ref16] je typickým a nejrozšířenějším otevřeným formátem pro distribuci rastrových geodat. Tento formát umožňuje uložit nejen rastrová data, ale také všechny typy gridových dat. Informace o souřadnicovém systému, souřadnicovém umístění a další popisné informace jsou uloženy přímo v hlavičce souboru. Při uložení dat do tohoto formátu nedochází při vhodné volbě komprese k nevratné ztrátě informace.

Formát JPEG [ref17] je kompresní formát určený k uložení rastrových souborů. Jeho výhodou je úspora místa a tudíž i menší nároky na datový přenos. Jeho nevýhodou je to, že komprese je ztrátová – formát tedy není vhodný pro použití v GIS, neboť data jsou nenávratně poškozena a jsou tak pro další zpracování nepoužitelná. Na druhou stranu, pokud obrázek JPEG je použit pouze jako podkladová vrstva (např. u leteckých snímků, u kterých se nepředpokládá žádné jiné využití), lze tak snížit nároky na datový tok. Soubory ve formátu JPEG jsou nejčastěji výsledkem volání WMS služby. Pokud jsou soubory šířeny samostatně, je potřeba je opatřit metadatovým souborem se souřadnicovým umístěním (koncovka .jpw).

Formát PNG [ref76] byl vytvářen jako nástupce formátů JPEG a GIF kvůli softwarovým patentům, které jsou v nich použité. Způsob komprese nepoškozuje ostré hrany. Z tohoto důvodu se tento formát využívá pro topografické podkladové mapy v prohlížečích službách WMS a WMTS. Komprese fotografií s množstvím gradientů již není tak efektivní. Pro uložení geodat je tento formát opět nevhodný, kvůli limitu barevné škály, částečně ztrátové kompresi dat a omezení na 3 barevné kanály + průhlednost. Soubory ve formátu PNG jsou nejčastěji výsledkem volání WMS služby. Pokud jsou soubory šířeny samostatně, je potřeba je opatřit metadatovým souborem se souřadnicovým umístěním (koncovka .pnw).

Formát GIF měl své využití v minulosti hlavně mezi webovými mapovými aplikacemi. GIF disponuje omezenou barevnou škálou, pro geodata je nevhodný (nejedná-li se o data binární nebo s rozsahem hodnot 0-255). Z tohoto důvodu byl GIF nahrazen zmíněným modernějším formátem PNG. Pro GIF platí to samé, co pro soubory JPEG a PNG, je nejčastěji výsledkem volání WMS služby. Pokud jsou soubory v tomto formátu šířeny samostatně, musí u nich být přítomen metadatový soubor se souřadnicovým umístěním (koncovka .gfw).

Ostatní formáty pro uložení rastrových dat nedosáhly takového rozšíření v oblasti GIS jako formát GeoTIFF. Řada z nich je proprietárních a jsou používány často pouze oborově. Jedná se například o formáty MrSID, BMP, ArcSDE Raster a další.

OGC GML [ref19] jako otevřený standard je perspektivním formátem pro přenos vektorových dat. Jedná se o jednosouborový textový formát založený na značkovacím jazyce XML, je proto interpretovatelný i bez speciálního software. Kromě standardizace na úrovni OGC je definován technickou normou ISO

19136. Vzhledem k tomu je podporován většinou moderních GIS nástrojů. GML je také předepsaný technickými dokumenty INSPIRE a výchozím formátem služby WFS.

GML se používá jako univerzální formát pro data, která mohou mít i komplikovanější stromovou strukturu. Díky tomu, že je postaven na XML, je jeho strojové zpracování jednoduché i běžnými systémy, například pomocí transformace XSLT.

Formát CityGML [ref79] je formát založený na XML, určený k reprezentaci souborů městských objektů ve 3D. Pomocí tohoto formátu je možné reprezentovat třídy, jejich vazby a vztahy relevantních topografických objektů ve městech a respektovat přitom jejich geometrické, topologické a sémantické vlastnosti. Pomocí tohoto formátu lze dosáhnout také určité generalizace, popsat hierarchické vazby mezi objekty, agregace atp.

CityGML je odvozený od formátu GML verze 3 a je vhodný zejména při tvorbě analýz nad daty v městském prostředí, simulací, správě budov a podobně.

OGC KML je určen především pro vizualizaci jednotlivých prvků geodat. Formát byl původně vyvinut firmou Google a je postavený na jazyce XML. Data v souborech KML, na rozdíl od GML, umožňují použít pouze souřadnicový systém WGS 84.

KML podporují produkty firmy Google, ale i řada služeb a programů třetích stran. Bývá často podporován moderními GPS přijímači. V minulosti býval nasazován ve webových mapových aplikacích, protože je v porovnání s GML menší a obsahuje zmíněnou informaci o vizualizaci jednotlivých prvků geodat. Ačkoliv byl v době před cca 3 lety tento formát populární, dnes je často nahrazován formátem GeoJSON.

Populárními formáty se v poslední době stávají formáty odvozené z formátu JSON, především GeoJSON a TopoJSON. Dalším odvozeným formátem od JSON je formát firma Mabox [ref82] používaným zejména pro předgenerované vektorové soubory uložené ve formě dlaždic pro rychlejší přenos dat.

Formáty JSON mají své uplatnění především mezi webovými technologiemi. Oproti formátům odvozených z XML (GML, KML) mají kratší zápis, což je výhodné při přenosech v prostředí Internetu. Stejně jako při využití formátů odvozených z XML, je i zde možné zabezpečit správnost struktury dat to pomocí schémat.

JSON je velice přívětivý k netyповým programovacím jazykům, je srozumitelný prostým lidským okem. Souřadnicový systém není v těchto formátech jak specifikovat, předpokládá se, že se jedná o WGS 84. Data lze libovolným způsobem zanořovat a větvit.

GeoJSON [ref68] je využíván u webových služeb pro svůj malý objem a jednoduchost. Je méně náročný na zpracování, což je vhodné zejména u webových prohlížečů. U uživatelů mimo svět GIS je oblíbený, protože jeho strukturu je možné rychle pochopit a připravit vlastní parser.

TopoJSON je druhým formátem odvozeným z formátu JSON, který ale zatím nenabyl takové popularity jako Geo-JSON. Hlavním úkolem formátu TopoJSON je minimalizace datového toku mezi webovým serverem a klientem. Formát je částečně ztrátový, neboť souřadnice bodů a lomových bodů jsou zapisovány v relativní poloze od daného počátku a v celých číslech (ztrácí se přesnost). K úspoře datové velikosti vede také fakt, že např. hranice polygonů jsou uloženy pro dvě sousedící plochy pouze jednou (formát je tedy topologický).

Formát TopoJSON je velice slibný, v budoucnu se očekává jeho rozšíření. V tuto chvíli naráží zejména na nedostatečnou podporu v softwarech. Není ani vhodný jako obecný formát pro výměnu dat mezi systémy. Formát byl navržen s ohledem na optimalizaci aplikací ve webovém prostředí a tam má taky své místo.

Geodatabáze Spatialite je postavená na souborové Open Source databázi SQLite. SQLite je přítomna v řadě zařízení či programech, interně ji využívá např. prohlížeč Firefox. Spatialite je její prostorové

rozšíření, podobně jako PostGIS pro databázi PostgreSQL. SpatialLite umožňuje uložit a pracovat s geodaty v prostředí SQL databáze, která je ovšem uložena v jednom jednoduše přenositelném souboru.

SpatialLite je vhodný formát na lokální uložení dat, ale v praxi se pro výměnu dat příliš nepoužívá.

Formát SHP je v praxi již dlouhou dobu nejpoužívanějším formátem pro výměnu vektorových geodat [ref18]. Bohužel je tento formát v dnešní době již poněkud omezující, zejména z důvodů zmíněných níže. Stále je ale používán pro menší datové soubory a jednoduché datové sady bez komplikovaných vazeb mezi objekty a tabulkami, protože je to formát jednoduchý a poskytuje jistotu kompatibility mezi různými softwarovými platformami.

Mezi slabá místa formátu patří zejména to, že data nejsou uložena v jednom souboru, ale hned ve trojici (shp+shx+dbf). Různé softwarové produkty si navíc přidávají vlastní metadatové soubory, které nejsou součástí specifikace tohoto formátu¹. Názvy atributů jsou omezeny pouze na deset znaků. Data neobsahují informaci o znakové sadě, což vede k problémům při automatické konverzi dat a používání na více operačních systémech. Velikost souborů je maximálně 2GB. Neumožňuje ukládat topologické informace o vzájemných vztazích mezi prvky geodat. Každý soubor SHP umožňuje ukládat pouze jeden typ geometrie (bod, linie, polygon) a neumožňuje uložit stromovou strukturu dat. Moderním nástupcem výše zmiňovaných rastrových, ale především vektorových formátů je standard OGC GeoPackage [ref39]. Tento formát umožňuje uložit libovolná vektorová data spolu s daty rastrovými, ať už ve formě dlaždic, nebo souborů ve formátu GeoTIFF do prostředí databáze SQLite. Poskytuje tak jednoduché rozhraní jazyka SQL pro práci s daty². Vektorová data jsou uložena dle specifikace OGC Simple Features for SQL [ref40]. Maximální velikost databázového souboru je 140 TB, což je pro praktické použití většinou dostačující. Data v jedné datové vrstvě, tedy databázové tabulce, mohou mít různé typy geometrií. Řada GIS nástrojů již podporu pro OGC GeoPackage nabízí, včetně Open Source knihovny GDAL od verze 1.11 či proprietárního prostředí Esri ArcGIS od verze 10.2.1.

OGC GeoPackage se zatím v praxi příliš nepoužívá. Nicméně vzhledem k tomu, že se jedná o standard OGC umožňující práci s velice komplexními datovými strukturami, jsme toho názoru, že by se tento formát měl pro otevřená geodata využívat a to i přesto, že podpora tohoto formátu není v běžných programech mimo svět GIS příliš rozšířena.

Distribuce otevřených geodat

Na způsob distribuce libovolných dat má vliv mnoho faktorů, zejména životní cyklus poskytovaných dat a typ uživatele, který je bude využívat.

S ohledem na životní cyklus dat je třeba rozlišovat mezi statickými daty a těmi, které se průběžně mění (dynamická data). Příkladem statických dat jsou výstupy analýz a data popisující konkrétní stav. Data, která se v čase mění můžeme potom dále dělit na dva základní okruhy. Do prvního náleží taková data, která popisují v reálném čase se měnící jev, to může být například znečištění, demografická data atd. Druhým typem jsou data, která nepopisují proměnlivý jev, ale jsou průběžně nebo nárazově aktualizována. Takovými daty může být například digitální model reliéfu.

Typ uživatele je druhým z faktorů, který je vhodné mít na paměti při volbě způsobu distribuce geodat. S určitou mírou zjednodušení lze konstatovat, že čím jsou data komplexnější, tím obtížnější je jejich uchopení na straně příjemce. Základním způsobem distribuce dat pro uživatele by měly být webové služby OGC. V tomto případě získává uživatel vždy nejaktuálnější data. Náročnějším způsobem získávání dat je tzv. vytěžování (harvesting) poskytovaných dat a budování databáze na vlastním hardware. U dat, která jsou průběžně aktualizována, je v těchto případech nutné umožnit získávání stavových dat (tj. dat platných k určitému datu) a změnových souborů.

¹ Shoda napříč programy panuje alespoň na souboru s příponou .prj, který obsahuje informace o souřadnicovém systému.

² Formát OGC GeoPackage byl navržen primárně jako výměnný formát. Nejedná se o alternativu ke geodatabázím jako je např. SpatialLite, které nabízejí podporu pro pokročilé prostorové SQL dotazy (*spatial SQL queries*).

Specifickou oblastí u poskytování otevřených dat je poskytování dat agregovaných (znenpřesněných nebo bez některých atributů). Obvyklým důvodem agregace³ bývají citlivé údaje (osobní údaje, data vlastněná třetími stranami) obsažené v datech, které se nesmí poskytovat třetím osobám. V tomto případě je nutné geodata od těchto údajů očistit před samotnou distribucí.

Výdejní systém, má-li být efektivní a funkční, musí kopírovat charakter dat, nad kterými je postaven. Nelze jej vyvíjet nezávisle na datech, které má vydávat.

Obsahem kapitoly jsou vhodné způsoby distribuce otevřených geodat, zejména pomocí webových služeb OGC a publikačního standardu Atom. Nakonec bude uvedena alternativní možnost publikace geodat pomocí služby GitHub.

2.2.3 Implementační pravidla INSPIRE

Jedním z osvědčených způsobů distribuce geodat v Evropské unii je využití prohlížečích, stahovacích služeb a vyhledávacích služeb podle směrnice INSPIRE, která se také opírá o standardy konsorcia OGC. O tom, že směrnici INSPIRE, resp. technické dokumenty s ní svázané, lze považovat za "best practice" svědčí i to, že podobné postupy se prosazují i jinde ve světě, například na Novém Zélandu [ref46]. Popis implementace jednotlivých částí směrnice je obsažen v tzv. implementačních pravidlech. Na publikaci vektorových a rastrových dat se vztahuje technický průvodce [ref28].

Technický průvodce pro implementaci *Stahovací služby INSPIRE* se dotýká právě problematiky velkých datových sad. Nabízí dvě možnosti implementace této služby:

- Předgenerované soubory s datovou sadou a jejich distribuce prostřednictvím dokumentu ve formátu Atom – ovšem bez možnosti jejich dotazování či výběru částí dat prostřednictvím serveru.
- Webové služby OGC WFS a WCS (tak zvané *datové sady s přímým přístupem*). Ty rozšiřují možnosti předgenerovaných datových sad o možnost filtrovat požadovaná data již na straně serveru.

V obou případech je k dispozici tzv. Get Download Service Metadata Request. V prvním případě seznam odkazů ve formátu Atom (viz kapitola *Předgenerované soubory a formát Atom*), v druhém případě pomocí WFS nebo WCS GetCapabilities.

Implementační pravidla směrnice INSPIRE definují také požadavky na dostupnost služeb, jejich kapacitu a rychlost odezvy. Praxe ukazuje, že požadavky definované v technických specifikacích INSPIRE jsou velice vágní, nedostatečně specifikované a v praxi dokonce podhodnocené (např. požadovaná dostupnost služby 99% znamená, že služba může být nedostupná 3.65 dne v roce!).

2.2.4 Otevřené webové služby - OGC OWS

Základním způsobem distribuce otevřených geodat je využití otevřených webových standardů OGC Open Web Services (OWS). Nejpoužívanějšími službami jsou OGC Web Map Service, Web Feature Service a Web Coverage Service. Existují však i jiné standardy, mající opodstatnění v některých případech použití. Standardy OGC jsou postaveny nejčastěji na komunikaci mezi serverem a klientem prostřednictvím zpráv ve formátu XML. Tyto standardy jsou podporovány ve většině používaných GIS softwarech. OGC služby jsou použité i v technických normách směrnice INSPIRE.

Při distribuci pomocí webových služeb získává uživatel vždy nejaktuálnější data. Nevýhodou tohoto způsobu je zátěž na straně infrastruktury poskytovatele, kterou není možné vždy předvídat, konzument navíc očekává garanci jejich dostupnosti. Zátěž serverů je potřeba průběžně sledovat a adekvátně na ni

³ Agregace je seskupení vybrané části určitých entit za účelem vytvoření nové entity – seskupení datových prvků do větších skupin. Tím dochází k odstranění některých detailů z dat. Agregace se může použít např. pro znenpřesnění nebo anonymizaci citlivých dat.

reagovat. V tomto směru může být cestou pro distribuci otevřených geodat využití cloudového řešení na pronajatých sdílených serverech, kde je výkon dynamicky zvyšován podle potřeby a cena potom odpovídá využití. K tomu je však potřeba překonat určitou psychologickou bariéru, jelikož data a infrastruktura zdánlivě nejsou pod kontrolou jako v případě, že poskytovatel použije vlastní řešení.

V této části jsou zmíněny pouze nejčastěji používané standardy, které pokrývají většinu případů použití:

- OGC Web Map Service
- OGC Web Map Tiled Service
- OGC Web Feature Service
- OGC Web Coverage Service
- OGC Sensor Observation Service

Služba OGC WMS [ref20] je standard, pomocí kterého může klient požádat o mapový obraz ve formě rastrového souboru. Server jej na základě klientských požadavků vytvoří a klientovi odešle. Klient specifikuje obsah obrázku (zobrazené vrstvy), souřadnicový systém, hraniční souřadnice, velikost, formát obrázku a další možné detaily. Server odešle pouze rastrový obraz dat, nikoli vlastní data. To lze s výhodou využít pro případ poskytnutí dat pouze k nahlédnutí z důvodu, že poskytovatel nechce nebo nemůže zpřístupnit zdrojová data. Standardním formátem výstupu je obrázek ve formátech PNG nebo JPEG podle charakteru dat.

Služba OGC WMTS [ref23] je používána v případě, kdy se data v čase příliš nemění (například letecké snímky, obecně podkladové mapy) a lze pro ně na straně serveru připravit předgenerované dlaždice (tiles - obrázky o pravidelné velikosti, většinou 256x256 pixelů) do vyrovnávací paměti pro určitá měřítká a v určitém rozsahu (cache). Tyto dlaždice pak lze zpřístupnit podle standardu WMTS (nebo i WMS). Výhodou je rychlé odbavení příchozího požadavku a nižší zátěž IT infrastruktury. Nevýhodou je, že dlaždice musí být omezeny pro určitá měřítká. Obsah je statický (v čase se nemění, datové vrstvy vykreslené v obrázku jsou stále stejné). Takto vytvořenou databázi dlaždic je potřeba udržovat, pravidelně aktualizovat a mít pro ni dostatečně velkou diskovou kapacitu. Standardním formátem výstupu je obrázek ve formátech PNG nebo JPEG podle charakteru dat.

Jako vhodná sada měřítek spolu s výchozím “počátkem” dlaždic se ukazuje řada dlouhodobě používaná servery ČÚZK, který pro souřadnicový systém S-JTSK (EPSG:5514, dříve EPSG:2065 či ESRI:102067) vyvinul vlastní řadu měřítek [ref24]. Pro globální souřadnicové systémy (jako je např. “Spherical Mercator” EPSG:3857) se doporučuje používat měřítkovou řadu vyvinutou firmou Google.

Služba OGC WFS [ref21] slouží k distribuci vektorových dat. Standard WFS 2.0.0 umožňuje také spouštět některé analytické operace přímo na serveru, jsou-li na něm podporovány. WFS dále podporuje filtrování požadovaných prvků geodat (vzhled jevů, *features*), není tak potřeba stahovat celou datovou sadu. Pro větší objemy dat je možné použít možnost stránkování odpovědi, tj. nemusí být stahována všechna data najednou v jedné odpovědi. Pomocí WFS může server vrátit data v libovolném formátu, který podporují knihovny pracující na pozadí (např. SHP, GeoJSON, ...), standardní bývá formát GML.

Služba OGC WCS [ref22] slouží k distribuci rastrových dat. Tento standard je vhodný zejména tam, kde chceme uživatelům nabídnout ke stažení velká rastrová data, která mohou být i multispektrální, či mohou obsahovat více rozměrů. Standardním formátem výstupních dat bývá GeoTIFF.

Služba OGC SOS [ref72] je vhodná pro zpřístupnění měření ze senzorů a senzorových sítí, stejně jako pro jejich popis. Sensory většinou publikují několik měření k danému místu a v daném čase. Poloha senzoru může být statická, ale může se i v čase měnit. Sensory mohou měřit různé veličiny a v různých časových úsecích. Nemění se celý dataset, ale získáváme časovou řadu měření.

2.2.5 Předgenerované soubory a formát Atom

Pro datové sady větších objemů je vhodné předgenerovat jejich obsah do cílových vektorových formátů a postavit kolem nich architekturu, která v nich umožní efektivně vyhledávat. Jedním z vhodných nástrojů je např. formát Atom [ref25]. U dat, která jsou průběžně aktualizována, je nutné umožnit jak získávání stavových dat (tj. dat platných k určitému datu), tak změnu formou předgenerovaných změnových souborů. Režim poskytování dat je vhodné nastavit s ohledem na objem změn. Toto řešení často vede ke snížení zátěže na infrastrukturu poskytovatele.

Tento způsob se blíží populárnímu a velice jednoduchému přístupu "vystavit soubory na FTP server". To se s formátem Atom nevylučuje - Atom slouží pouze jako metadatový dokument, ze kterého lze rychle vyčíst referenci k cílovým souborům.

Atom je v principu XML dokument, který obsahuje odkazy na dostupné datové sady nebo soubory a jejich základní metadata. Tento formát je využíván i v dalších technologických standardech, jako je například OGC OWS Context [ref38].

Soubor ve formátu Atom je webový standard pro publikování syndikovaného obsahu. Syndikovaný obsah je takový obsah, který na webu již může být publikován, souborem Atom se mu ale zpětně přidají vybraná metadata a tím se zjednodušeně popíše pro automatické zpracování. Atom má nahradit starší (proprietární a stále populární formát RSS) a je původně určen pro webové stránky. Nicméně jeho využití pro data se nabízí.

Soubor Atom obsahuje hlavičku, která identifikuje vlastní zdroj a autora a seznam "záznamů", také s jednoznačnou identifikací a vlastním obsahem nebo odkazem na tento obsah.

Příklad formátu atom je uveden v *Příloha G: Formát Atom*.

2.2.6 Služby GitHub

Služba GitHub [ref41] je webové rozhraní k systému pro správu verzí Git, který byl původně napsán za účelem správy a údržby zdrojového kódu jádra operačního systému GNU/Linux. Od roku 2014 je možné do této služby nahrávat i geografická data v některých z formátů GeoJSON a TopoJSON. GitHub soubory uložené v těchto formátech umí přímo vizualizovat. Podle různých údajů lze usoudit, že limit pro velikost vstupního souboru, má-li být zobrazen v mapové prohlížečce, je v současnosti cca 4.5 MB, záleží ale také na struktuře vstupního souboru [ref42]. U jednodušších struktur může být limit až 10 MB (maximální velikost souboru na serverech GitHub je cca 100 MB). Pokud je datový soubor příliš veliký, není zobrazen. Jeho praktickou dostupnost to ale nijak neovlivní.

Takto jednoduše publikovaná data lze stáhnout opět v jednom z podporovaných formátů. Výhoda tohoto přístupu je mimo jiné v tom, že poskytovateli dat zcela odpadá starost o IT infrastrukturu. O tu se stará třetí strana - v tomto případě GitHub. Uživatelé navíc získají efektivní nástroj pro verzování dat v čase. Pokud by byla služba GitHub v budoucnu uzavřena anebo by se změnila výrazně její obchodní politika, nejednalo by se o tak zásadní problém. Systém Git je decentralizovaný, každý uživatel má u sebe lokální kopii celé datové sady včetně veškeré historie. Vzhledem k tomu, že je systém pro správu verzí Git vyvíjen jako Open Source, tak by bylo možné případný přechod na jinou formu distribuce ze služby GitHub realizovat bez větších problémů.

Do prostředí GitHub lze nahrát i dlaždicovaná rastrová data a odkazovat se na ně formou zápisu identifikátoru URL podle standardu OGC TMS. Podle zkušeností uživatelů se jeví tato služba jako dostatečně rychlá.

Se službou GitHub již experimentují některé menší obce a samosprávy (např. [ref43], [ref44]), ale i větší města (např. Chicago [ref57]). Tento přístup k publikování geodat je vhodnější pro menší města

bez vlastního IT oddělení. Nicméně některé koncepty tohoto přístupu (správa verzí, distribuce, náhled, atd.) jsou aplikovatelné i na tuto případovou studii.

2.3 Verzování dat a časové řady

Geografická data nejsou již delší dobu omezena pouze na dvoudimenzionální prostor (2D). Data jsou často trojdimenzionální (3D a to jak gridová - volumes, tak vektorová). Mohou být ale i n-dimenzionální (v případě pásem družicových snímků). V případě časoprostorových dat je dalším rozměrem, který je potřeba zohlednit, čas. Potom mluvíme o 4D datech.

Na časovou složku v datech se můžeme dívat minimálně ze dvou pohledů: Datová sada může obsahovat "časovou řadu" nějakého fenoménu (např. vývoj teploty na daném území, pohyb senzorů v prostoru a čase, vývoj jejich hodnot) nebo změna verze celé datové sady (nové přesnější zaměření budov, stav k určitému datu a podobně).

Z hlediska distribuce a formátů dat se k oběma typům přistupuje stejně. Tam, kde je některá služba nebo formát vhodnější na jeden z typů časové složky je to zdůrazněno.

Standard WMS nabízí možnost, jak definovat další dimenze pro poskytovaná data. Nejčastější formou použití je právě čas, ale může to být např. nadmořská výška, teplota, atd. či případně i jejich kombinace. V metadatech služby lze uvést buď přesnou časovou specifikaci výčtem časových okamžiků nebo počáteční čas a velikost časového kroku mezi jednotlivými datovými vrstvami. Příklady jsou uvedeny v *Příloha B: OGC WMS – dimenze a časové řady*.

Standard WMTS navíc umožňuje definovat různé dimenze k předgenerovaným datovým sadám. Princip je podobný jako u zmíněného standardu WMS, příklady jsou uvedeny v *Příloha C: OGC WMTS – dimenze a časové řady*.

2.3.1 OGC Web Feature Service

Standard WFS nemá přímou podporu pro časovou dimenzi. Standard odkazuje na OGC Filter Encoding Specification (FES) [ref49], pomocí kterého lze filtrovat požadovaná data na základně požadavků ze strany klienta. Pomocí FES lze nastavit počáteční a koncový hraniční čas (startTime a endTime), mezi kterými klient požaduje stáhnout data. Verzovat lze také pomocí vlastních klíčových slov (např. číslem revize "1.2.3" nebo datem "2014-01-20" a podobně).

Z uvedeného vyplývá, že WFS slouží jako rozhraní k datové sadě, která obsahuje data v různých časových intervalech. Na data vztahená k určitému časovému okamžiku se lze dotazovat právě pomocí filtru dle standardu FES 2.0.

2.3.2 OGC Web Coverage Service

OGC WCS podporuje ve své nejnovější verzi specifikace [ref50] časový rozsah požadovaných dat jako jeden z možných rozměrů. Syntaxe pro definici času sleduje stejně jako u výše zmíněných služeb technickou normu ISO 8601. Příklad je uveden v *Příloha D: OGC WCS 2.0.0 – dimenze a časové řady*.

Podle ústního sdělení editora standardu OGC WCS Petra Baumana, se momentálně v rámci organizace OGC téma času zásadním způsobem mění, neboť se začínají zohledňovat různé kalendáře (historické, i používané v různých kulturách či technických společnostech) a další s touto problematikou související komplikace. Viditelné je to zejména na tom, že ve starších verzích standardů býval definován parametr TIME explicitně jako vstupní parametr. U nových verzích standardů se čas mění v jeden z rozměrů

dat. Stejně jako stávající rozměry mají své zobrazení a souřadnicový systém, musí mít i čas společnou referenci.

Verzování a časové řady u souborových formátů a jejich distribuce

2.3.3 OGC GeoPackage

Formát OGC GeoPackage [ref39] je postavený na souborové databázi SQLite (viz kapitola *geopackage-ref*), což umožňuje v porovnání se stávajícími souborovými formáty pokročilejší funkce pro dotazování a manipulaci s daty pomocí jazyka SQL. Lze využít standardních datových typů TIME a DATETIME jako atributu daného geoprvcu. Další důležitou vlastností je metadatová tabulka `gpkg_content`, obsahující mimo jiné informace `last_change` (datový typ DATETIME) pro jednotlivé tabulky (datové vrstvy). Dále existuje metadatová tabulka `gpkg_metadata`, obsahující vlastnost `timestamp`, kterou lze využít na označení aktuálnosti libovolné jednotky v souboru – buď celé databáze, jednotlivé tabulky či geoprvcu, tj. záznamu v tabulce.

2.3.4 Verzování systémem Git

Git je systém na správu verzí, nejčastěji textových souborů, viz kapitola *Služby GitHub*. To znamená, že pomocí Gitu lze udržovat přehled o souborech, o tom, kdo je měnil a jaké změny provedl. Případné konfliktní změny lze řešit poměrně komfortně, lze se “vracet v čase”, získat stav souboru k určité revizi nebo časovému okamžiku. Soubor s daty by měl být v Gitu uložen ideálně v textové podobě (GML, GeoJSON, ...). Binární formáty lze technicky vzato spravovat v prostředí Git také, potom ale nelze využít specializované verzovací nástroje.

2.3.5 Poskytování změnových souborů (RÚIAN best practice)

ČÚZK zavedl pro distribuci dat Registru Územní Identifikace, Adres a Nemovitostí (RÚIAN) systém měsíční aktualizace stavových dat s denními dávkami změnových souborů. Tento systém plně pokrývá jak potřeby uživatele, který potřebuje jednorázově získat podkladová data, tak uživatele, který potřebuje udržovat aktuální obraz celé databáze, aniž by byl nucen stahovat velké objemy dat po síti. Možnost získat seznam přírůstků od libovolného data zvyšuje na straně uživatele pružnost procesu aktualizace dat. *Datové sady jsou nabízeny v různě obsáhlých verzích, v některých případech je dokonce možné volit generalizované hranice. Data jsou nabízena buď pro celé území České republiky, anebo po jednotlivých obcích. To umožňuje při poměrně malé zátěži na straně serveru efektivně obsloužit velké množství klientů. Práce s aktualizací dat se přesouvá ze strany serveru ke klientům.*

V jedné věci se však RÚIAN nechová ideálně: Jednotlivé stavové a změnové soubory mají sice pevnou a strojově předvídatelnou strukturu, chybí jim však centrální strojově zpracovatelný zdroj. Tím by mohl být například zmiňovaný formát Atom (viz *Předgenerované soubory a formát Atom*). Podle ústního sdělení bude Atom v nejbližší době doplněn ⁴.

⁴ Ústní sdělení, konference "Inspirujme se ...", 2014

2.4 3D Data

3D data obsahují kromě svého umístění v prostoru i informaci o "hloubce". To se týká jak rastrových tak vektorových dat.

Nejtypičtějším příkladem 3D rastrových dat bývá digitální model reliéfu. V tomto případě se ale nejedná o plnohodnotná 3D data. V rastrové matici je pouze uložena výška povrchu, ale už ne informace o tom, co se děje "pod ní". Hovoříme tak o 2.5D datech.

Samozřejmě je možné uložit i plnohodnotná 3D rastrová data (tzv. volumes), obsahující např. informaci o půdním profilu, srážkovou mapu a podobně. Většina specializovaných GIS má pro podobná data vlastní formát. Pro technologicky neutrální distribuci prostorových dat však můžeme využít např. formát GeoTIFF, a jednotlivé vrstvy uložit jako "pásma" rastrového snímku.

Prakticky všechny formáty vektorových dat dnes umožňují uložení souřadnice z k lomovým bodům. Některé formáty obsahují i speciální 3D vektorové objekty (ekvivalent polygonu face, či ekvivalent 2D centroidu pro 3D objekt kernel).

Pro distribuci otevřených geografických dat ve 3D by pro většinu aplikací měl dostačovat běžný formát (GeoPackage, SHP, GeoJSON, GML, ...).

Pro speciální aplikace je vhodné zvážit, zda by bylo možné data distribuovat ve formátu CityGML (viz *citygml*). Tento formát umožňuje popsat nejen geometrický tvar tělesa, ale i vztahy mezi objekty (bloky budov, čtvrtě), vnitřní strukturu budov a podobně.

2.5 Metadata

Veškerá publikovaná geodata a na ně navazující webové služby je potřeba opatřit příslušnými metadaty. Metadata jsou strukturovaná data o datech. Metadata popisují data a služby ve strojově zpracovatelném formátu tak, aby bylo možné v nich automaticky vyhledávat a to i na základě jejich relevance a aktuálnosti. Pro metadata existuje množství standardů a doporučení, ale zdaleka ne všechny jsou vhodné pro oblast geodat.

Vlastní metadata mohou (měly by) mít jak vlastní datové sady (kdo je vytvořil, kdy, s jakou přesností, co obsahují, z jaké oblasti data jsou atd.), tak i webové služby tyto datové sady publikující (kdo provozuje danou službu, jaké datové sady služba publikuje, atd.). Metadata by měla být dostupná přes rozhraní webové služby OGC Catalog Service for Web (CSW) [ref37]. Zároveň doporučujeme tuto službu otestovat na dostupném software (Esri ArcGIS, QGIS a další) tak, aby byla ověřena její praktická funkčnost a dostupnost na různých platformách.

V současné době postupně do oblasti GIS a geodat pronikají vznikající standardy pro obecná otevřená data a to se týká i metadat. Otevřená provázaná data (open linked data) mají vlastní metadatové standardy, které jsou již v souladu s INSPIRE mapovatelné tak, aby bylo v těchto datových souborech možné vyhledávat pomocí CSW a obráceně, linkovaná geodata je možné publikovat na portálech s otevřenými daty.

V současné době je pro pořizování a uchovávání metadat v geodatové doméně klíčová mezinárodní technická norma ISO 19115 [ref32]. Tuto normu navíc vyžaduje i evropská směrnice INSPIRE ve svém nařízení komise o metadatech [ref33]. Vlastní technickou implementací této normy se zabývají implementační pravidla směrnice INSPIRE pro metadata [ref34]. Vlastní fyzické uložení metadat geografické datové sady nebo služby je definováno navazující technickou normou ISO 19139 (XML) [ref35]. Obecně lze říci, že je vhodné držet se metadatového profilu České republiky [ref36], i když to v první fázi za vysloveně nutné nepovažujeme.

2.6 Souřadnicové systémy

Systém souřadnic je soustava základních údajů (referenčních bodů, přímek nebo křivek), umožňující určovat souřadnice polohy objektu ve zvolené vztažné soustavě. Protože převod tvaru Země na plochu papíru (dnes monitoru počítače) je vždy provázen určitou nepřesností, existuje množství systémů, které v daném místě na Zemi poskytují známé a popsitelné zkrácené-nepřesnost (nebo globální systémy, používané pro zobrazení celé Země).

Poznámka: Protože existuje množství způsobů, jak popsat konkrétní systém a velké množství souřadnicových systémů, bývá v oboru zvykem, že se pro souřadnicové systémy používá databáze EPSG (European Petroleum Survey Group). Tu lze stáhnout ze stránek <http://www.epsg-registry.org/> nebo využívat některou ze služeb nad touto databází postavenou, např. <http://epsg.io>.

Geografické datové sady jsou v České republice vedeny především v souřadnicovém systému S-JTSK (EPSG 5514 ⁵). Pro vojenské mapové podklady se v minulosti používal souřadnicový systém S-42 (EPSG 3835 ⁶). Vzhledem k zániku Varšavské smlouvy a pozdějšímu přistoupení k NATO se začal místo souřadnicového systému S-42 používat systém UTM/WGS-84 (zóny 33 - EPSG 32633 ⁷ a 34 - EPSG 32634 ⁸). Evropská směrnice INSPIRE [ref26] zejména pak ve specifikaci věnované souřadnicovým systémům ([ref29], str. VII) a pro měřítka větší než 1:500 000 mezi podporované systémy přidává ETRS89-TM (EPSG 3035 ⁹). Praxe si vynutila použití souřadnicového systému Spherical Mercator (EPSG 3857 ¹⁰), zavedeného firmou Google pro jejich mapové produkty.

Poznámka: Dříve používané zápisy S-JTSK, jako EPSG 2065 či ESRI/ESPG 102067, vznikly díky tomu, že v databázi EPSG nebyl přítomný kód pro S-JTSK s "otočenými osami" (a zápornými hodnotami souřadnic), tzv. "S-JTSK/Krovak East North". To dnes již není potřeba a všechny systémy by měly nadále používat EPSG 5514.

Pro stahovací služby se přikláníme k publikování datových sad v jejich původních souřadnicových systémech, což je v praxi většinou S-JTSK. Navíc všechny relevantní desktopové GIS programy jsou schopny transformovat geodata do uživatelem požadovaných cílových souřadnicových systémů online. Pokud se přeci jen ukáže, že je potřeba poskytnout některým klientům možnost stahovat data v jiném souřadnicovém systému, doporučujeme zprovoznit transformační souřadnicovou službu podle specifikace INSPIRE [ref30].

V každém případě je potřebné zajistit, aby distribuovaná data měla korektně nastaveny definice souřadnicových systémů. V případě S-JTSK je nutné, aby informace o souřadnicovém systému obsahovala parametry pro transformaci mezi referenčním Besselovým elipsoidem a elipsoidem WGS-84, transformační parametry anebo grid. Jinak může dojít k nepřesnosti při transformaci až v řádu několika desítek metrů. Více informací k tomuto tématu lze najít například na Portálu FreeGIS [ref31].

Vedle S-JTSK doporučujeme nabízet data v souřadnicovém systému WGS84 (EPSG 4326 ¹¹). Zejména zahraniční uživatelé či uživatelé kombinující data z různých datových zdrojů tento souřadnicový systém využijí. Kromě toho se používá v navigacích a GPS zařízeních.

U prohlížečích služeb je vhodné umožnit zobrazení dat v souřadnicovém systému Spherical Mercator, využívaný např. firmou Google ve svých mapových produktech, projektem OpenStreetMap a nebo ma-

⁵ <http://epsg.io/5514>

⁶ <http://epsg.io/3835>

⁷ <http://epsg.io/32633>

⁸ <http://epsg.io/32634>

⁹ <http://epsg.io/3035>

¹⁰ <http://epsg.io/3857>

¹¹ <http://epsg.io/4326>

pami Bing. To uživatelům umožní využívat tyto zobrazovací služby v kombinaci s jinými podklady. Tento systém nesl kdysi neoficiální označení 900913, nyní je již zastaralý a v EPSG je označen kódem 3857.

Obecně lze říci, že na běžných platformách webových serverů je přidání dalšího souřadnicového systému otázkou minimálního zásahu do konfigurace, takže lze podporu pro další systémy přidávat i na požádání.

Ve starších verzích standardů OGC se předpokládalo, že pořadí souřadnic v požadavku (např. parametru BBOX u WMS) nebo při odpovědi (např. GML publikované serverem WFS) je vždy ve formátu X,Y.

V nových verzích standardů (WMS 1.3.0, WFS 2.0.0, atd.) je explicitně zdůrazněno, že záleží na předpisu daného souřadnicového systému – pořadí os tedy může být X,Y ale i Y,X. To platí zejména pro souřadnicové systémy WGS-84 (EPSG 4326) a ETRS-89 (EPSG 3035). U S-JTSK (EPSG 5514) se tento fakt v praxi nezohledňuje.

Zdáleka ne všechny serverové ale i klientské implementace standardů jsou schopny pořadí souřadnic korektně zohlednit [ref54], což je dobré mít na paměti.

2.7 Závěrečná doporučení

Pro účely vymezené v začátku tohoto oddílu se jeví jako vhodná licence Creative Commons BY-SA 4.0 [ref47], případně Open Data Commons Attribution License (ODC-BY) [ref48]. Výhodou první je její obecná známost (i napříč veřejnou správou), druhá je lépe přizpůsobena pro využití v oblasti geodat.

Třetí možností je vytvořit licenci na míru danému projektu, a to případně i odvozením ze dvou zmíněných výše. V takovém případě je ale nezbytná spolupráce s expertem na autorské právo.

2.7.1 Formát souborů

Nelze jednoduše doporučit jeden či dva formáty vhodné pro všechny uživatele a datové sady. Vždy je potřeba zvážit charakter dat a převládající způsob jejich použití.

Pro předgenerované soubory vektorových dat doporučujeme, v dlouhodobém horizontu, formát OGC GeoPackage. V krátkodobém horizontu lze použít i formát SHP nebo GML, z toho důvodu, že formát GeoPackage není zatím příliš rozšířen.

Pro publikování formou prohlížečích webových služeb (OGC WMS, WMTS) je vhodné volit v závislosti na charakteru dat formáty PNG a JPEG.

V případě stahovacích služeb doporučujeme pro vektorová data formát GML a pro rastrová data potom podle jejich charakteru GeoTIFF či JPEG.

2.7.2 Distribuce otevřených geodat

Jako primární doporučujeme využít standardy OGC OWS, zejména WMS, WFS a WCS.

Kde to z důvodu velikosti datových sad nebo pro technická omezení na straně poskytovatele není možné, doporučujeme předgenerovat datové soubory ve vhodném datovém formátu a poskytnout soubor ve formátu Atom s odkazy na takto vytvořené soubory. Podobně jako se k tomu kloní implementační pravidla INSPIRE.

Pro datové sady, které se *mění v čase* a jsou příliš velké na to, aby se s každou změnou vydávala aktualizovaná verze celé sady, je vhodné publikovat jednou v pravidelných intervalech stavová data a současně k nim poskytovat v kratších časových intervalech změnové soubory. Toto řešení může výrazně snížit

zatížení IT infrastruktury, neboť uživatelé nemusí vždy stahovat celou datovou sadu ve formě stavových dat, ale pouze menší změnové soubory, které si sami aplikují na kopii datové sady tak, aby ji měli co možná nejaktuálnější. Více k tomuto tématu v kapitole *Předgenerované soubory a formát Atom*. Více o časových řadách v části *Verzování dat a časové řady*.

Souřadnicové systémy

2.7.3 Prohlížeční a stahovací služby

Pro prohlížeční službu OGC WMTS je výhodné nabízet předgenerované dlaždice minimálně pro souřadnicové systémy EPSG 3857 (Spherical Mercator) a to ve schématu, ve kterém je nabízí např. mapy firmy Google, ale i projekt OpenStreetMap [ref52] (nebo služby ČÚZK) a pro souřadnicový systém S-JTSK (EPSG 5514) ve schématu používaném serverem ČÚZK [ref53].

Transformace mezi souřadnicovými systémy bývá u současných software poměrně rychlá. Služby by měly nabízet data především v původním souřadnicovém systému - převážně S-JTSK (EPSG 5514). Vzhledem k určité exotičnosti tohoto systému je vhodné navíc podporovat minimálně ETRS-89 (EPSG 3035) a WGS 84 (EPSG 4326). Spherical Mercator (EPSG 3857) není v tomto případě nezbytný, většina webových aplikací transformuje vektorová data z WGS84 automaticky, stejně tak většina desktopových nástrojů.

2.7.4 Předgenerované vektorové a rastrové soubory

Vzhledem k plánovaným větším objemům dat doporučujeme publikovat datové soubory v původním souřadnicovém systému (S-JTSK, EPSG 5514) a případně WGS 84 (EPSG 4326). Implementační pravidla směrnice INSPIRE doporučují poskytnout transformační službu, která umožní na straně serveru transformovat data přímo do cílového systému.

Otevírání geografických dat

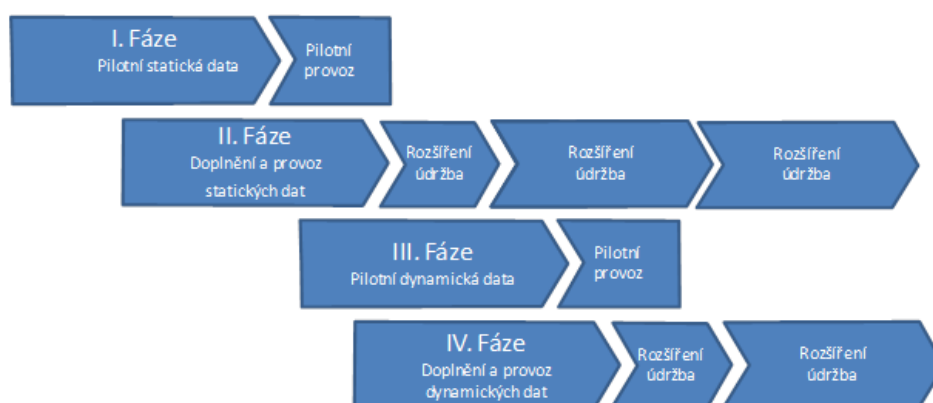
3.1 Doporučený postup otevírání geografických dat

K otevření geodat je vhodné postupovat v jednoznačně oddělených fázích, které na sebe navazují, což umožní využít získané zkušenosti a aplikovat požadavky uživatelů, které při provozu otevřených dat získáme.

Jednotlivé odbory městských samospráv často vytvářejí velké množství kvalitních dat, která jsou potenciálně atraktivní pro širokou škálu uživatelů. Na druhou stranu zdroje pro otevírání dat jsou omezené, proto je potřeba pečlivě vybírat, co se bude uvolňovat. Rizikem je vyčerpání energie na datové sady, které nebudou uživateli příliš využívány a naopak nepokrytí těch, které by byly pro veřejnost více atraktivní.

Pro cílový stav doporučujeme, aby otevřená data byla publikována jako *webové služby OGC OWS*. Tam, kde se to ukáže jako problematické (např. z důvodu přílišné zátěže na stávající technickou infrastrukturu), doporučujeme poskytování statických datových sad, tj. těch, které se vygenerují, připraví na server včetně vyhledávací infrastruktury s využitím *formátu Atom*. V druhém kroku, s využitím zkušeností při otevírání statických dat, je vhodné přistoupit k otevírání dat dynamických, měnících s v čase. Každý krok je rozdělen na *pilotní fázi*, ve které jsou otevřeny vhodně vybrané datové sady, u kterých předpokládáme velký zájem uživatelů, a *fázi implementační*, ve které jsou aplikovány zkušenosti z pilotní fáze do přípravy rutinního provozu a systémového otevírání dat. Otevírání dalších datových sad pokračuje postupně tak, aby byla neustále zabezpečena kvalita poskytovaných dat a služeb.

Doporučený postup otevírání dat je znázorněn na následujícím schématu:



Obrázek 3.1: Doporučený postup otevírání dat

Příprava datových sad pro zveřejnění by měla zahrnovat následující kroky [ref70]:

1. Identifikace datových sad pro otevření,
2. kontrola kvality vybraných datových sad - kompletnost, aktuálnost a přesnost, geometrická správnost,
3. přidání/kontrola metadat a jejich údržba,
4. volba a přidání licenčních podmínek,
5. fyzické otevření dat publikováním stahovacích/prohlížečích služeb.

Pro identifikaci datových sad je vhodné zvolit kombinaci z následujících možností:

1. dotazníkové šetření v komunitě uživatelů
2. analýza požadavků stávajících uživatelů
3. analýza nákladů na zveřejnění a porovnání s náklady na stávající systém distribuce
4. analýza složitosti zveřejnění (legislativní omezení, čištění dat, ...)
5. odhad potřeb/pozorování uživatelů

Důležité je nejenom jednorázově provést průzkum a odhad potřeb uživatelů, ale také umožnit uživatelům zpětnou vazbu směrem k poskytovateli, minimálně pomocí formuláře na webových stránkách.

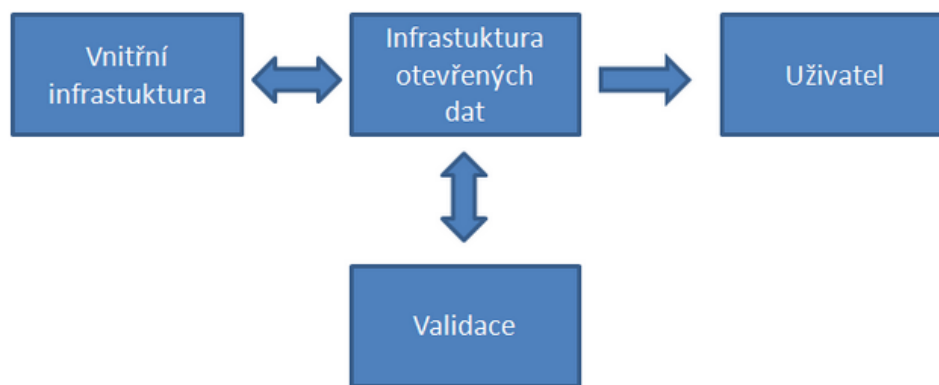
Otevírání dat je proces, který nikdy nebude možné označit za ukončený. Proto není cílem otevřít všechny datové sady, které jsou momentálně k dispozici, ale pracovat na jejich otevírání postupně.

Velice důležitým krokem se jeví popularizace s cílem získat maximální spektrum potenciálních uživatelů. Zveřejnění v aktualitách portálu města je potřeba periodicky opakovat pro všechny nové uživatele, minimálně v intervalu dvakrát ročně. Manažerskou veřejnost je nutné oslovit propagací na konferencích a workshopech. Odbornou veřejnost a vývojáře potom pomocí hackfestů, hackathonů, cílených workshopů, cenami pro nejlepší produkt a pod.

3.2 Technické řešení infrastruktury otevřených dat

Při návrhu a budování infrastruktury otevřených dat je vhodné dodržovat několik základních pravidel, která zabezpečí maximální efektivitu, rozšiřitelnost a odolnost celého řešení.

- Vnitřní infrastrukturu je nutné logicky, ne nutně fyzicky, oddělit od infrastruktury pro poskytování otevřených dat.
- Data jsou publikována v souladu s platnými standardy. Je možné, že stávající softwarové vybavení nebude tuto podmínku splňovat. Softwarové řešení může být hybridní (Open Source a proprietární), každopádně takové, aby plnilo daný účel.
- Pokud to licenční podmínky umožní, je vhodné v maximální míře využít stávajících technologií tak, aby se zbytečně nerozšiřovalo spektrum provozovaného software a tím i složitost celého řešení. To znamená, že optimální je vnitřní infrastrukturu geografických dat a infrastrukturu dat otevřených provozovat na stejné platformě, doplněné o případné nutné další technologie tam, kde fyzicky stávající technologie nedostačuje.
- Klíčovým prvkem infrastruktury je validace dat před jejich publikováním validátorem příslušného standardu. Validaci samotnou je nutné svěřit jiné části týmu, případně externistům, aby byly nalezeny i nepopsané vlastnosti, sdílené pouze vnitřními pracovníky.
- Provoz dynamických služeb je vhodné delegovat na externí infrastrukturu, která umožní zabezpečit její škálovatelnost a dostupnost podle aktuálního zatížení.



Obrázek 3.2: Technická infrastruktura při otevírání dat

Nasazené řešení je vhodné testovat na co možná nejširší škále běžně používaných nástrojů z oblasti desktopových GIS aplikací, geodatabází, katalogů a pod. Výběr referenčního softwaru je vždy ovlivněn osobní preferencí administrátora, objektivními vlastnostmi, znalostí určitého technologického okruhu.

3.2.1 Doporučení k validaci

Publikovaná data v souborových formátech i webové služby by mělo jít snadno přidat jako vrstvy do tzv. projektů (workspace) v desktopových aplikacích jako je Esri ArcGIS, GeoMedia, QGIS, případně Topol a další. Pro ověření správnosti definice souřadnicového systému poskytovaných dat lze použít transformační nástroje Open Source knihovny GDAL [ref59]. Dále je pomocí této knihovny vhodné ověřit konzistenci a validitu dat včetně nastavení kódování češtiny v poskytovaných datech. To je možné zpětným importem a porovnáním s původními daty. Knihovna GDAL je taktéž využívána systémem Esri ArcGIS, lze ji tedy považovat za vhodnou referenční knihovnu. Pro úplnost dodáváme, že za ekvivalentní řešení ke knihovně GDAL pro prostředí jazyka Java lze považovat knihovnu GeoTools [ref63].

Provázání mezi metadatovým katalogem a samotnou službou je vhodné testovat pomocí zásuvného modulu QGIS MetaSearch [ref58], který umožňuje v katalogu přidat službu jako vrstvu do projektu. Dále by mělo být možné ověřit validitu služeb (vyhledávacích ale i prohlížečích a stahovacích) a dat na národním geoportálu INSPIRE [ref61].

Pro automatické testování validity a dostupnosti webových služeb můžeme s výhodou využít knihovnu OWSLib [ref62], kterou lze spouštět periodicky na serveru a provádět tak pravidelnou kontrolu.

Pro ověření funkčnosti a dostupnosti webových služeb, ale i předgenerovaných rastrových dlaždic prostřednictvím služby OGC WMTS, je vhodné využít i webové aplikace, např. pomocí knihoven Leaflet [ref64] nebo OpenLayers [ref65]. Více v *Příloha A: Open Source software pro jednotlivé úlohy*.

U poskytovaných geodat, kde předpokládáme migraci na straně uživatele do geodatabáze, jakou jsou například objemnější data v relační nebo stromové struktuře, případně soubory se změnovými větami, je na místě tuto migraci otestovat. Pakliže se zaměříme na Open Source řešení, tak se jako nejvhodnější referenční geodatabáze jeví PostGIS [ref60]. Zde je třeba počítat s tím, že postup nebude triviální a není od věci jej publikovat, například na stránkách IPR Praha.

Příloha A: Open Source software pro jednotlivé úlohy

Pro jednotlivé úlohy spojené s publikováním otevřených geografických dat (převod do správného formátu, publikace pomocí služeb nebo statických souborů, validace, ...), je potřeba software. Vždy doporučujeme využít maximálně stávající technické (software i hardware) vybavení v organizaci dostupné, může se ale stát, že se ukáže, že stávající platforma je z nějakého důvodu (kapacitní, licenční, zastaralost, ...) nevhodná.

Zde představené programy jsou mezi uživateli rozšířené a představují obvyklé řešení při budování geografického informačního systému postaveného na Open Source software. Vývojáři Open Source software se primárně snaží o maximální implementaci stávajících standardů (často se jedná o referenční implementace) a maximální operabilitu. Také proto jsou tyto systémy vhodné i v hybridních systémech kombinujících software s otevřeným i uzavřeným zdrojovým kódem.

Je iluzorní automaticky předpokládat, že použitím Open Source software musí nutně dojít v dané organizaci k významnému ušetření finančních prostředků. Především záleží na kvalitě pracovníků, kteří se mohou, ale nemusejí spolehnout na vlastní síly při údržbě celého systému a od toho se následně odvíjí i cena celého řešení. V konečné sumě ale odpadají nutné náklady na pořízení software a na udržování licencí. Open Source lze nasadit okamžitě podle potřeb, lze s ním okamžitě experimentovat a vybudovat potřebné řešení, bez nutnosti vypořádávat předem licenční otázky.

Poznámka: Seznam zdaleka není kompletní a pro další alternativy doporučujeme navštívit stránku <http://osgeo.org> (Open Source Geospatial Foundation), která většinu programů s otevřeným zdrojovým kódem zaštiťuje.

A.1 Geodatabáze PostGIS

PostGIS [ref60] je nadstavba Open Source objektově-relační databáze PostgreSQL určená pro práci s geografickými daty. PostGIS lze ve světě Open Source bez nadsázky považovat za vedoucího hráče na poli geodatabází a jako přímou alternativu k proprietárnímu Oracle Spatial. Z hlediska správy geodat PostGIS nabízí mnohem více než jen úložiště dat, ale také mocné dotazovací prostředí, včetně podpory pro prostorové funkce, integraci topologické správy vektorových dat a společnou správu pro rastrová data. V posledních verzích PostgreSQL se objevuje podpora pro formáty jako JSON a jeho binární obdoba BSON, což umožňuje vybudovat NOSQL databázi a velice efektivně “kešovat” strukturovaná data pro jejich rychlý a efektivní výdej. To může v případě prozíravě navrženého řešení, při současném zachování vnitřní integrity dat, podstatně snížit požadavky na IT infrastrukturu.

Další zjevnou výhodou nasazení PostGIS je jeho provázanost s celým ekosystémem Open Source, ale i proprietárního softwaru. Pokud tedy budujeme systém, který má být “bezešvý”, je využití PostGIS ideálním řešením.

A.2 Mapový server MapServer

Program MapServer [ref66] je tradiční Open Source software pro distribuci geodat. Podporuje relevantní standardy OGC, díky napojení na knihovnu GDAL umožňuje export do požadovaných formátů v rámci webových služeb, podporuje časovou složku u služby OGC WMS. Díky nadstavbě MapCache umožňuje generovat dlaždice a nabídnout je prostřednictvím OGC WMTS. Při porovnání jednotlivých dostupných implementací OGC WMS se ukázalo, že se jedná o nejrychlejší software ze všech zúčastněných [ref67].

A.3 Mapový server GeoServer

Program GeoServer [ref80] je tradiční Open Source software pro distribuci geodat. Primárně se zaměřuje na implementaci standardů OGC, obsahuje nadstavby na tvorby rastrových dlaždic, geoprocessingové služby a další.

GeoServer je naprogramovaný v jazyce Java a množství práce v něm obstarává knihovna GeoTools. Na rozdíl od MapServeru disponuje webovým uživatelským rozhraním, které je pro většinu uživatelů relativně rychle pochopitelné.

A.4 Knihovna GDAL

Knihovna GDAL [ref59] je bezesporu základním kamenem Open Source GIS ekosystému. Podporuje čtení a zápis více než 100 nejrůznějších rastrových a více než 80 vektorových formátů včetně proprietárních formátů jako Esri File Geodatabase či specifických českých formátů VFK a VFR. Díky integrované knihovně Proj.4 podporuje transformaci dat i do těch nejexotičtějších souřadnicových systémů. Výhodná je kombinace této knihovny s programovacím jazykem Python, která nabízí uživateli velmi mocný a flexibilní nástroj pro práci s geodaty. Právě skripty v jazyce Python by mohly potencionálně tvořit jádro řešení pro otevírání geodat implementované v Open Source frameworku. Kromě generování souborů či integraci s mapovým serverem MapServer je možné knihovnu GDAL použít i pro implementaci sady testů validity a konzistence publikovaných geodat.

A.5 PyCWS

PyCWS [ref69] je serverová implementace standardu OGC CWS napsaná v jazyce Python. PyCWS umožňuje publikovat a vyhledávat v metadatech geografických datových sad. Pomocí tohoto serverového řešení lze postavit v rámci infrastruktury geodat (SDI) katalog metadat odpovídající v současnosti platné technické normě ISO, evropské směrnici INSPIRE a souvisejícím národním profilům. PyCWS je certifikovaný OGC software.

A.6 GeoNetwork

GeoNetwork [ref81] je serverová implementace standardu OGC CWS napsaná v jazyce Java umožňující publikovat a vyhledávat v metadatech geografických datových sad. Pomocí tohoto serverového řešení lze postavit v rámci infrastruktury geodat (SDI) katalog metadat odpovídající v současnosti platné technické normě ISO, evropské směrnici INSPIRE a souvisejícím národním profilům. GeoNetwork je velice oblíbený právě v organizacích implementujících směrnici INSPIRE.

Příloha B: OGC WMS – dimenze a časové řady

B.1 Příklad OGC WMS Capabilities response s časovou řadou

```

...
<Layer>
<Name>time</Name>
<Title>Časová vrstva - Příklad</Title>
...
<Dimension name="time" units="ISO8601" default="2003-10-17">
1996-01-01/2003-10-17/P1D
</Dimension>
...
</Layer>
...

```

B.2 Různý způsob vyjádření obsahu elementu Dimension

Syntaxe	Význam
hodnota	Samostatná jedna hodnota
hodnota1,hodnota2,hodnota3,...	Seznam více hodnot
min/max/krok	Interval definovaný minimální a maximální hodnotou a krokem
min1/max1/krok1,min2/max2/krok2,...	Více různých intervalů

B.3 WMS GetMap požadavek s parametrem TIME

```
...&TIME=1996-01-02&...
```

Příloha C: OGC WMTS – dimenze a časové řady

C.1 Příklad OGC WMTS Capabilities response s časovou řadou

```
...
<Layer>
<Name>time</Name>
<Title>Časová vrstva - Příklad</Title>
...
<Dimension>
  <ows:Identifier>time</ows:Identifier>
  <UOM>ISO8601</UOM>
  <Default>2013-05-29</Default>
  <Current>true</Current>
  <Value>2012-05-08/2013-05-29/P1D</Value>
</Dimension>
...
</Layer>
...
```

C.2 WMTS požadavek s parametrem TIME

C.2.1 Kódování Key-Value-Pairs (KVP)

```
...&TIME=1996-01-02&...
```

C.2.2 Kódování REST

Šablona

```
http://foo/wmts/{Product}/{Style}/{Time}/{TileMatrixSet}/{TileMatrix}/
{TileRow}/{TileCol}.jpg
```

Konkrétní příklad

<http://foo/wmts/ortofoto/default/2012-07-09/EPSSG5514/6/13/36.jpg>

Příloha D: OGC WCS 2.0.0 – dimenze a časové řady

Níže uvedené příklady předpokládají, že čas je vedle souřadnic X,Y veden jako další rozměr dat. Čas je zadáván v souladu s technickou normou ISO 8601.

D.1 Příklad požadavku WCS s časovým omezením - Key Value Pairs

```
http://www.myserver.org:port/path?service=WCS&version=2.0&request=GetCoverage
&coverageId=C0002
&subset=long(100,200)
&subset=lat(50,60)
&subset=days("2009-11-06T23:20:52Z","2009-11-06T23:20:52Z")
```

D.2 Příklad požadavku WCS s časovým omezením - XML POST

```
<gmlcov:GridCoverage ... gml:id="C0001">
  <gml:boundedBy>
    <gml:Envelope srsName="http://www.opengis.net/def/crs-compound?
      1=http://www.opengis.net/def/crs/EPSG/0/4326&
      2=http://www.opengis.net/def/crs/OGC/0/Days"
      axisLabels="Lat Long Time" uomLabels="deg deg day" srsDimension="3">
      <gml:lowerCorner>1 1 2012-03-10</gml:lowerCorner>
      <gml:upperCorner>3 10 2012-03-10 </gml:upperCorner>
    </gml:Envelope>
  </gml:boundedBy>
  ...
</gmlcov:GridCoverage>
```

Příloha E: Geometrická validita geoprvků

U dat poskytovaných veřejnou správou se obvykle předpokládá určitá míra jejich návaznosti. Proto je nezbytné, aby data byla po geometrické stránce validní, a to i proto, že existuje velice úzká souvislost mezi validitou dat a jejich jednoznačností. Jinak řečeno, validní data nemohou být po geometrické stránce dezinterpretována a budou v různých softwarech zobrazena vždy stejně. Za geometricky validní data považujeme taková data, která jsou v souladu se specifikací OGC Simple Features [ref51].

Obzvláště u prostorových dat exportovaných ze systémů, nesloužících primárně jako geografické informační systémy, jako je například CAD, můžeme nezděka najít chybné geometrie. Nejčastější chybou bývají nesprávně "zaplochované" polygony, zdvojené hranice mezi dvěma polygony, překřížené hranice sousedících ploch a další chyby v topologii. Výsledkem bývá například chybný výpočet plochy, který neodpovídá skutečné ploše popisovaného jevu.

K tomuto problému můžeme přistupovat z různých pohledů. V ideálním případě pořizujeme data způsobem, který vylučuje jejich uložení v případě, že nejsou po geometrické stránce v pořádku. Na druhou stranu to není vždy technicky možné. Často pracujeme s přebíranými daty, kde by bylo pro jejich vyčištění do nich nutné neúměrně zasáhnout. To však není vždy možné. Platí zde přímá úměra, čím jsou data závaznější, důležitější, tím obtížnější bývá jejich čištění. V takovém případě není vždy možné poskytovat data dokonale validní a korektní. V takové situaci je třeba spolu s daty udržovat aparát pro popis chyb a u jednotlivých záznamů uvádět, zda jsou plně validní. Příkladem tohoto přístupu mohou být data o ZCHÚ a NATURA poskytovaná AOPK. Vhodnou formou se zdá být záznam o chybách v metadatech.

Příloha F: Kódování češtiny

Pro zapsání české diakritiky do tabulky znaků existuje z historických důvodů několik možných způsobů. Do dnešní doby přetrvávají na některých systémech kódování CP852, Windows-1250 či ISO-8859-2. Díky tomu dříve často docházelo (a někdy stále dochází) k nekoretnímu zápisu a čtení z databází, kdy české znaky bývají zkomolené. Uživatel se tak může spolehnout pouze na svůj úsudek či náhodu, aby se mu znaky zobrazily korektně.

Problémy s nedostatečně velkou tabulkou znaků řeší kódování UTF-8, do kterého lze uložit znaky všech známých jazyků, čímž by se měly problémy s kódováním vyřešit.

Je proto nezbytné, aby kódování znaků dat publikovaných jako otevřená data, bylo stejné u všech souborů a to *UTF-8*.

Příloha G: Formát Atom

Atom (Atom Syndication Format) je webový standard pro publikování syndikovaného obsahu. Je nástupcem populárního formátu RSS.

Samotný kanál a jednotlivé jeho položky musí mít uvedený název, jedinečný identifikátor (URI) a datum poslední změny. Data (text, XML či binární data) mohou být přímo součástí kanálu, to ovšem pro geodata není z pochopitelných důvodů vhodné a proto je lepší uvést odkaz (*link*) na datovou sadu.

Níže uvádíme hypotetický příklad kanálu ve formátu Atom, který obsahuje 3 položky, a to vrstevnice s přesností 1, 2 a 5 metrů.

```
<?xml version="1.0" encoding="utf-8"?>
<feed xmlns="http://www.w3.org/2005/Atom">

  <title>Otevřená data IPR Praha</title>
  <link href="http://geoportalpraha.cz"/>
  <link href="http://geoportalpraha.cz/opendata/feed" rel="self"/>
  <updated>2015-01-01T12:00:00+0100</updated>
  <author>
    <name>Josef Novák</name>
    <email>josef.novak@iprpraha.cz</email>
  </author>
  <id>ipr:data.urm.cz:2B7AB7CF-A6B0-4AB8-BB2B-229AC168605C</id>

  <entry>
    <title>Vrstevnice 1 m</title>
    <link href="http://geoportalpraha.cz/data/vrstevnice_1m.gpkg"/>
    <id>CZ-URM-MAP.MAP_vrstevnice_1m_1</id>
    <updated>2010-12-16</updated>
    <summary>Vrstevnice jsou vygenerované z Digitálního modelu terénu 2010.
      Vrstevnicový plán odpovídá obsahem a přesností úrovni map 1:5000.
      Přesnost cca 1m.
    </summary>
  </entry>

  <entry>
    <title>Vrstevnice 2 m</title>
    <link href="http://geoportalpraha.cz/data/vrstevnice_2m.gpkg"/>
    <id>CZ-URM-MAP.MAP_vrstevnice_2m_1</id>
    <updated>2010-12-16</updated>
    <summary>Vrstevnice jsou vygenerované z Digitálního modelu terénu 2010.
      Vrstevnicový plán odpovídá obsahem a přesností úrovni map 1:5000.
      Přesnost cca 2m.
    </summary>
  </entry>
</feed>
```

```
</entry>

<entry>
  <title>Vrstevnice 5 m</title>
  <link href="http://geoportalpraha.cz/data/vrstevnice_5m.gpkg"/>
  <id>CZ-URM-MAP.MAP_vrstevnice_5m_1</id>
  <updated>2010-12-16</updated>
  <summary>Vrstevnice jsou vygenerované z Digitálního modelu terénu 2010.
    Vrstevnicový plán odpovídá obsahem a přesností úrovni map 1:5000.
    Přesnost cca 5m.
  </summary>
</entry>

<!-- ... Další záznamy ... -->
</feed>
```

- [ref01] Mráček, Jakub a kol.: *Jak otevírat data?*, Fond Otakara Motejla, <http://www.otevrenadata.cz/res/data/001/003498.pdf> [cit. 2014-09-06]
- [ref02] Fond Otakara Motejla: *První pražský GeoData Hackathon*, <http://www.nasstat.cz/aktualni-deni/1-prazsky-geodata-hackathon/> [cit. 2014-09-06]
- [ref03] GISMentors: *Výsledky pražského GeoData Hackathonu (13. - 15. 6. 2014 IPR Praha)*, <http://www.gismentors.eu/geodatahackathon2014/> [cit. 2014-09-06]
- [ref04] Otevřená společnost o.p.s.: *Mapa kriminality*, <http://mapakriminality.cz/> [cit. 2014-09-06]
- [ref05] JH Partners, s.r.o.: *Váš majetek.cz*, <http://www.vasmajetek.cz/> [cit. 2014-09-06]
- [ref06] Rozpočet veřejně o.p.: *Rozpočet obce*, <http://www.rozpocetobce.cz/> [cit. 2014-09-06]
- [ref07] Ministerstvo financí: *ÚFIS: Monitoring hospodaření obcí*, <http://www.info.mfcr.cz/ufis/>, [cit. 2014-09-06]
- [ref08] Zajíček, Petr.: *Vliv příjmové politiky města na jeho výdaje*. (Diplomová práce) Praha, VŠE, 2014, https://www.vse.cz/vskp/37084_vliv_prijmove_politiky_mesta_na_jeho_vydaje
- [ref09] Pípová, Tereza.: *Výdaje na veřejnou správu měst*. (Diplomová práce) Praha, VŠE, 2014, https://www.vse.cz/vskp/37094_vydaje_na_veřejnou_spravu_mest
- [ref10] IHned.cz: *Jak daleko dojedete pražskou hromadnou dopravou*, <http://ihned.cz/mhd>, [cit. 2014-09-06]
- [ref11] Centrum aplikované ekonomie, o.s.: *Politické finance*, <http://www.politickefinance.cz/> [cit. 2014-09-06]
- [ref12] ČUZK: *Registr územní identifikace, adres a nemovitostí*, <http://ruian.cz> [cit. 2014-09-06]
- [ref13] Plan4Business: *What's the plan?* <http://www.whatstheplan.eu/> [cit. 2014-09-06]
- [ref14] Jarmič, L., Záhumenská, V.: *Územní plánování stručně a jasně aneb kdy a jak se efektivně zapojit*, Ekologický právní servis, Brno, <http://www.eps.cz/o-nas/publikace> [cit. 2014-09-06]
- [ref15] Tim Berners-Lee: *Linked Data*, <http://www.w3.org/DesignIssues/LinkedData.html> [cit. 2014-09-06]
- [ref16] OSGeo: *GeoTIFF*, <http://trac.osgeo.org/geotiff/> [cit. 2014-09-06]
- [ref17] Joint Photographic Experts Group: *JPEG*, <http://jpeg.org> [cit. 2014-09-06]

- [ref18] Esri: *ESRI Shapefile Technical Description*, ESRI, 1998, <http://www.esri.com/library/whitepapers/pdfs/shapefile.pdf> [cit. 2014-09-06]
- [ref19] Portele, C. (editor): *OpenGIS Geography Markup Language (GML) Encoding Standard*, OGC 07-036, Open Geospatial Consortium, Inc., 2007, <http://www.opengeospatial.org/standards/gml> [cit. 2014-09-07]
- [ref20] Beaujardiere, J. (editor): *OpenGIS Web Map Server Implementation Specification*, OGC 06-042, Open Geospatial Consortium, Inc., 2006, <http://www.opengeospatial.org/standards/wms> [cit. 2014-09-07]
- [ref21] Vretanos, P. (editor): *OpenGIS Web Feature Service 2.0 Interface Standard*, OGC 09-025r1, Open Geospatial Consortium, Inc., 2010, <http://www.opengeospatial.org/standards/wfs> [cit. 2014-09-07]
- [ref22] Baumann, P. (editor): *OGC® WCS 2.0 Interface Standard- Core: Corrigendum*, OGC 09-110r4, Open Geospatial Consortium, Inc., 2012, <http://www.opengeospatial.org/standards/wcs> [cit. 2014-09-07]
- [ref23] Masó, J. a kol.: *OpenGIS Web Map Tile Service Implementation Standard*, OGC 07-057r7, Open Geospatial Consortium, Inc. 2010, <http://www.opengeospatial.org/standards/wmts> [cit. 2014-09-07]
- [ref24] ČÚZK: *Prohlížečí služba WMTS - Základní mapy ČR*, http://geoportal.cuzk.cz/WMTS_ZM/WMTService.aspx?service=WMTS&request=GetCapabilities [cit. 2014-09-07]
- [ref25] IETF AtomPub Working Group: *Atom Publishing Protocol*, <http://www.atomenabled.org/> [cit. 2014-09-07]
- [ref26] CENIA, česká informační agentura životního prostředí: *INSPIRE*, <http://inspire.gov.cz/o-inspire> [cit. 2014-09-07]
- [ref27] *Zákon 380/2009 Sb.*, http://inspire.gov.cz/sites/default/files/documents/380_2009.pdf [cit. 2014-09-07]
- [ref28] Initial Operating Capability Task Force for Network Services: *Technical Guidance for the implementation of INSPIRE Download Services*, version 3.1, Initial Operating Capability Task Force, 09-08-2013 <http://inspire.ec.europa.eu/index.cfm/pageid/5> [cit. 2014-09-08]
- [ref29] INSPIRE Thematic Working Group Coordinate Reference Systems & Geographical Grid Systems: *Data Specification on Coordinate Reference Systems – Technical Guidelines*, INSPIRE Thematic Working Group Coordinate Reference Systems & Geographical Grid Systems, 2014-04-17, D2.8.I.1_v3.2, http://inspire.jrc.ec.europa.eu/documents/Data_Specifications... [cit. 2014-09-08]
- [ref30] Network Services Drafting Team: *Draft Technical Guidance for INSPIRE Coordinate Transformation Services*, Network Services Drafting Team, 2010-03-15, http://inspire.jrc.ec.europa.eu/documents/Network_Services/... [cit. 2014-09-08]
- [ref31] Portál FreeGIS: *S-JTSK*, <http://freegis.fsv.cvut.cz/gwiki/S-JTSK> [cit. 2014-09-08]
- [ref32] *ISO 19115 Geographic Information - Metadata. International Organization for Standardization (ISO)*, Geneva, 2003
- [ref33] *NAŘÍZENÍ KOMISE (ES) č. 1205/2008*, kterým se provádí směrnice Evropského parlamentu a Rady 2007/2/ES týkající se metadat, Úřední věstník Evropské unie, 2008, <http://inspire.gov.cz/dokumenty/metadatech/208-narizeni-o-metadatech-cz> [cit. 2014-09-16]
- [ref34] Massimo Craglia (Editor), INSPIRE Metadata Implementing Rules: *Technical Guidelines based on EN ISO 19115 and EN ISO 19119*, verze 1.2, European Commission Joint Research Centre,

- Ispra, 2007-10-26, <http://inspire.gov.cz/dokumenty/metadata/213-pokyny-k-narizeni-o-metadatech> [cit. 2014-09-16]
- [ref35] *ISO 19119:2011 Geographic information – Services*, International Organization for Standardization (ISO), Geneva, 2003
- [ref36] Kafka Š., Jirásková L., Kondrová L., Řezník T. (editoři): *Metadatový profil ČR pro soubory prostorových dat*, sérií souborů prostorových dat a služeb založených na prostorových datech, verze 2.0, Technická pracovní skupina Metadata, <http://inspire.gov.cz/dokumenty/metadata/616-metadatovy-profil-cr-v20> [cit. 2014-09-17]
- [ref37] Douglas Nebert, Arliss Whiteside, Panagiotis (Peter) Vretanos (editoři): *OpenGIS® Catalogue Services Specification*, Version 2.0.2, číslo dokumentu OGC 07-006r1, Open Geospatial Consortium, 2007-02-23, <http://opengeospatial.org/standards/cat> [cit. 2014-09-17]
- [ref38] Roger Brackin, Pedro Gonçalves (editoři): *OGC OWS Context Conceptual Model*, Version: 1.0, číslo dokumentu OGC 12-080r2, Open Geospatial Consortium 2014-01-22, <http://www.opengeospatial.org/standards/owc>, [cit. 2014-09-17]
- [ref39] Paul Daisey (editor): *OGC® GeoPackage Encoding Standard*, Version: 1.0, číslo dokumentu OGC 12-128r10, Open Geospatial Consortium 2014-02-12, <http://www.opengeospatial.org/standards/geopackage> [cit. 2014-09-17]
- [ref40] John R. Herring (editor): *OpenGIS® Implementation Standard for Geographic information - Simple feature access - Part 2: SQL option*, Version: 1.2.1, číslo dokumentu OGC 06-104r4, Open Geospatial Consortium 2010-08-04, <http://www.opengeospatial.org/standards/sfs> [cit. 2014-09-17]
- [ref41] GitHub*, <http://github.com>, [cit. 2014-09-17]
- [ref42] Mike Skalnik: *View GeoJSON/TopoJSON Source*, <https://github.com/blog/1865-view-geojson-topojson-source> [cit. 2014-09-17]
- [ref43] Map of Atlanta region street network*, Atlanta Regional Commission <http://atlregional.github.io/fc-review/>, <https://github.com/aaronr/foss4g-2014/blob/master/smalltown/slides.md> [cit. 2014-09-18]
- [ref44] Aaron Racicot: *Small town GIS - Leveraging GitHub*, QGIS and community members to manage local data, FOSS4G 2014, Portland, <http://vimeo.com/106228441> [cit. 2014-09-18]
- [ref45] CKAN, <http://ckan.org/> [cit. 2014-09-21]
- [ref46] Brent Wood: *Adding value to Open Data using Open Source GIS*, FOSS4G 2014, Portland, <https://vimeo.com/106229154> [cit. 2014-09-21]
- [ref47] Creative Commons: *Attribution-Share-Alike 4.0 International*, <https://creativecommons.org/licenses/by-sa/4.0/> [cit. 2014-10-24]
- [ref48] Open Knowledge Foundation: *Open Data Commons Attribution License*, <http://opendatacommons.org/licenses/by/summary/> [cit. 2014-10-24]
- [ref49] Peter Vretanos (editor): *OpenGIS Filter Encoding 2.0 Encoding Standard*, OGC 09-026r1 and ISO 19143:2010(E), Open Geospatial Consortium, 2010-11-22, <http://www.opengeospatial.org/standards/filter> [cit. 2014-09-27]
- [ref50] Peter Baumann, Jinsongdi Yu (editors): *OGC® Web Coverage Service Interface Standard - Range Subsetting Extension*, OGC 12-040, Open Geospatial Consortium, <http://www.opengeospatial.org/standards/wcs> [cit. 2014-09-27]

- [ref51] John R. Herring (editor): *Implementation Standard for Geographic information - Simple feature access - Part 1: Common architecture*, OGC 06-103r4, Version: 1.2.1, Open Geospatial Consortium, <http://www.opengeospatial.org/standards/sfa> [cit. 2014-09-28]
- [ref52] OpenStreetMap: *Slippy map tilenames*, http://wiki.openstreetmap.org/wiki/Slippy_map_tilenames [cit. 2014-27-09]
- [ref53] Srovnal Pavel: *Prohlížeč služba WMTS - Základní mapy ČR*, Zeměměřický úřad, http://geoportal.cuzk.cz/WMTS_ZM/WMTService.aspx?service=WMTS&request=GetCapabilities [cit. 2014-09-27]
- [ref54] Jáchym Čepický: *Pořadí os souřadnicových systému ve WFS*, Les-ejk.cz, <http://les-ejk.cz/2013/03/poradi-os-souradnicovych-systemu-v-wfs/> [cit. 2014-09-27]
- [ref55] ČÚZK: *Struktura výměnného formátu informačního systému katastru nemovitostí České republiky*, [http://www.cuzk.cz/Katastr-nemovitosti/Poskytovani-udaju-z-KN/Vymenny-format-KN/Vymenny-format-ISKN-v-textovem-tvaru/Popis_VF_ISKN-v5_0-\(1\).aspx](http://www.cuzk.cz/Katastr-nemovitosti/Poskytovani-udaju-z-KN/Vymenny-format-KN/Vymenny-format-ISKN-v-textovem-tvaru/Popis_VF_ISKN-v5_0-(1).aspx) [cit. 2014-09-29]
- [ref56] Young Hahn: *Vector Tiles for MapBox Streets*, MapBox Blog, <https://www.mapbox.com/blog/vector-tiles/> [cit. 2014-09-29]
- [ref57] City of Chicago GitHub, <https://github.com/Chicago> [cit. 2014-09-29]
- [ref58] Tom Kralidis: *MetaSearch Catalogue Client (plugin)*, QGIS Python Plugins Repository, verze 0.3.1, <https://plugins.qgis.org/plugins/MetaSearch/> [cit. 2014-09-29]
- [ref59] GDAL. 2014. GDAL - Geospatial Data Abstraction Library: *Version 1.11.0*, Open Source Geospatial Foundation, <http://gdal.osgeo.org> [cit. 2014-09-29]
- [ref60] Kolektiv autorů: *PostGIS 2.0 Manual*, pro verzi PostGIS 2.0.7SVN, <http://postgis.net/docs/manual-2.0/> [cit. 2014-09-29]
- [ref61] Cenia: *Národní geoportál INSPIRE - Validace*, Cenia <http://geoportal.gov.cz/web/guest/validate/metadata/> [cit. 2014-09-29]
- [ref62] Tom Kralidis a kol.: *OWSLib*, verze 0.8.9, <http://geopython.github.io/OWSLib/> [cit. 2014-09-29]
- [ref63] Kolektiv autorů: *GeoTools*, verze 11.0, OSGeo.org, <http://www.geotools.org/> [cit. 2014-09-29]
- [ref64] Vladimír Agafonkin a kolektiv: *Leaflet*, verze 0.7.3, <http://leafletjs.com> [cit. 2014-09-29]
- [ref65] Kolektiv autorů: *OpenLayers*, verze 2.12, <http://openlayers.org> [cit. 2014-09-29]
- [ref66] Kolektiv autorů: *MapServer*, verze 6.4, <http://mapserver.org> [cit. 2014-09-29]
- [ref67] Kolektiv autorů: *Benchmarking WMS*, OSGeo.org, http://wiki.osgeo.org/wiki/Benchmarking_2011 [cit. 2014-09-29]
- [ref68] Kolektiv autorů: *Formát GeoJSON*, <http://geojson.org> [cit. 2014-09-29]
- [ref69] Kolektiv autorů: *pycsw*, verze 1.10, <http://pycsw.org> [cit. 2014-09-30]
- [ref70] Kolektiv autorů: *The Open Data Handbook*, Open Knowledge Foundation, 2012, <http://opendatahandbook.org> [cit. 2014-09-30]
- [ref71] Kolektiv autorů: *Directive 2003/98/EC of the European parliament and of the Council*, <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2003:345:0090:0096:EN:PDF> [cit. 2014-10-10]

- [ref72] Arne Bröring (editor): *OGC® Sensor Observation Service Interface Standard*, Open Geospatial Consortium, OGC 12-006, Version 2.0, <http://www.opengeospatial.org/standards/sos> [cit. 2014-10-08]
- [ref73] MZe: *LPIS - veřejný registr půdy*, <http://lpis.cz/> [cit. 2014-10-09]
- [ref74] WikiProject: *Rules for import (and use) of Czech free geodata to OSM project in Czech Republic*, http://wiki.openstreetmap.org/wiki/WikiProject_Czech_Republic/freemap [cit. 2014-10-09]
- [ref75] *London Datastore*, <http://data.london.gov.uk/> [cit. 2014-10-09]
- [ref76] Kolektiv autorů: *Portable Network Graphics (PNG) Specification (Second Edition)*, ISO/IEC 15948:2003, W3C Recommendation 10 November 2003, <http://www.w3.org/TR/PNG/> [cit. 2014-11-05]
- [ref77] *Geografická informace – Značkovací jazyk geografie (GML)*, Česká technická norma, ČSN EN ISO 19136,
- [ref78] Capgemini Consulting: *The Open Data Economy: Unlocking Economic Value by Opening Government and Public Data*, http://www.capgemini-consulting.com/resource-file-access/resource/pdf/opendata_pov_6feb.pdf [cit. 2014-11-22]
- [ref79] Gerhard Gröger (editor) a kol.: *OGC City Geography Markup Language (CityGML) Encoding Standard*, OGC 12-019, Verze 2.0.0, Open Geospatial Consortium, Inc. 2012, <http://www.opengeospatial.org/standards/citygml>, [cit. 2014-11-30]
- [ref80] Kolektiv autorů: *GeoServer*, <http://geoserver.org> [cit. 2014-12-01]
- [ref81] Kolektiv autorů: *GeoNetwork*, <http://geonetwork-opensource.org> [cit. 2014-12-01]
- [ref82] Springmeyer, Dane a kol.: *Mapbox Vector Tile Specification*, <https://github.com/mapbox/vector-tile-spec> [cit. 2015-01-02]

Symbols

ČSÚ

Český statistický úřad, 14

ČÚZK

Český úřad zeměměřický a katastrální, 14

Časové řady, 38, 39

Čeština, 45

Český úřad zeměměřický a katastrální

ČÚZK, 14

Český statistický úřad

ČSÚ, 14

3D data, 28

A

Atom, 24, 25

Formát Atom, 47

B

Berlín, 14

Best practices, 12

C

Chicago, 14

CityGML, 21

CKAN, 5

CSW, 29

D

Děčín, 14

Distribuce geodat, 23

E

EPSG, 29

Esri Shapefile

SHP, 21

F

Fivestar system

Pěti hvězdičkový systém, 8

Formát, 17

Formát Atom

Atom, 47

Formát souboru, viz Formát

Formáty, 20

FTP, 25

G

GDAL, 38

GeoNetwork, 38

GeoPackage, 21

GeoServer, 38

GeoTIFF

TIFF, 21

GIF, 21

GitHub, 26

GML, 14, 21

H

Hackathon, 12

I

Implementační pravidla INSPIRE, 24

INSPIRE, 24

ISO 19115, 29

ISO 19139, 29

J

JPEG, 21

K

KML, 21

L

Londýn, 13

LPIS, 15

M

MapServer, 37

Metadata, 29

Ministerstvo zemědělství

Mze, 15
Mze
Ministerstvo zemědělství, 15

O

OGR, 38
Opava, 14
Open Data
 Otevřená data, 8
Open Data Handbook, 5
Open source software, 35
OpenStreetMap, 5
Otevřená data
 Open Data, 8
Otevřené formáty, viz Formáty

P

Pěti hvězdičkový systém
 Fivestar system, 8
pair OGC CSW, 29
PNG, 21
PostGIS, 37
PyCSW, 38

R

RÚIAN, 14
Rastry, 21

S

S-42, 29
S-JTSK, 29
SHP
 Esri Shapefile, 21
Souřadnicové systémy, 29
SpatialLite, 21

T

Technické řešení, 34
TIFF
 GeoTIFF, 21
tripple: JSON
 GeoJSON
 TopoJSON, 21

U

UTF-8, 45
UTM, 29

V

Vídeň, 14
Validace, 35
Veřejná správa, 5

Vektory, 21
Verzování, 39
VFR, 14

W

WGS84, 29
WMS, 38
WMTS, 39