

TB0500MV003 Metodika pro publikování
prostorových informací ve formě otevřených
dat

Otakar Čerba
Štěpán Kafka
Jan Ježek
Dzmitry Kozhukh
Václav Čada
Karel Charvát
Stanislav Holý
Tomáš Mildorf

2016

Obsah

Seznam zkratk	6
Seznam příloh	10
Úvod	11
1 Slovník základních pojmů	14
2 Metodika (Návrhová část)	17
2.1 Licence a ochrana osobních údajů	17
2.1.1 Ochrana osobních údajů	17
2.1.2 Otevírání dat ve vztahu k zákonu o ochraně utajovaných skutečností.	18
2.1.3 Práva duševního vlastnictví a otevřené licence	20
2.2 Metadata	22
2.2.1 Shrnutí požadavků na metadata	22
2.2.2 Detailní popis	23
2.2.2.1 Klíčová slova	23
2.2.2.2 Aplikační schéma / struktura dat	24
2.2.2.3 Licence	26
2.2.2.4 Adresy	27
2.2.2.5 Organizace	28
2.2.2.6 Prostorový rozsah dat	28
2.2.2.7 Účel vytvoření	29
2.2.2.8 Modelování vazeb	30
2.2.3 Příklady	33
2.2.4 Postup při publikaci otevřených dat	33

2.2.4.1	Scénář 1 - Publikování přes Národní katalog INSPIRE	33
2.2.4.2	Scénář 2 - Publikování přes Národní katalog otevřených dat	33
2.3	Formáty prostorových dat	34
2.3.1	Přehled používaných formátů a software	34
2.3.2	Hlavní otevřené datové formáty pro prostorová data . .	36
2.3.2.1	Comma Separated Values (CSV)	36
2.3.2.2	Keyhole Markup Language - KML/KMZ . . .	37
2.3.2.3	Geography Markup Language - GML	37
2.3.2.4	ShapeFile SHP	38
2.3.2.5	GeoJSON	38
2.3.3	Rastrová data	39
2.3.4	Otevřené datové služby	39
2.4	Publikování datových modelů	40
2.4.1	Jednoduchý datový model	40
2.4.2	Středně složitý datový model	41
2.4.3	Komplexní datový model	41
2.5	Identifikátory	41
2.5.1	Registr osob (ROS)	42
2.5.2	Registr obyvatel (ROB)	43
2.5.3	Registr územní identifikace, adres a nemovitostí (RÚIAN)	43
2.5.4	Registr práv a povinností (RPP)	44
2.6	Vazby na externí datové zdroje	44
2.6.1	Identické a podobnostní vazby	45
2.6.1.1	owl:sameAs	45
2.6.1.2	rdfs:seeAlso	45
2.6.1.3	skos:closeMatch	46
2.6.1.4	skos:exactMatch	46
2.6.2	Topologické vazby	47
2.6.3	Heirarchické a mereologické vazby	49
3	Modelové aplikace	52
3.1	Metadata	52
3.1.1	MICKA	52
3.1.2	Katalogový klient na Národním geoportálu INSPIRE .	52
3.1.3	On-line převodník	55
3.2	Ontologie VF XML	55

3.3	Smart Points of Interest	58
3.4	Open Transport Map	62
3.4.1	Datový model OTM	64
3.4.2	Převod dat do RDF	65
3.5	Open Land Use Map	69
3.5.1	OLU datový model	71
3.6	Adresy	74
4	Závěr	77
	Použitá literatura	81

Seznam tabulek

1	Příklady formátů v GIS.	34
2	Slovník RCC8.	48
3	Ukázka OTM číselníku.	64

Seznam obrázků

1	Uživatelské rozhraní katalogu MICKA. Červené šipky označují možnost exportu GeoDCAT. Vlevo - feed, vpravo jednotlivý záznam.	53
2	Uživatelské rozhraní katalogového klienta. Červená šipka označuje možnost exportu jednoho metadatového záznamu do GeoDCAT.	54
3	Zkušební rozhraní převodníku.	55
4	Ukázka objektových relací Ontologie VF XML.	56
5	Ukázka anotačních vlastností Ontologie VF XML.	57
6	Vizualizace Ontologie VF XML.	59
7	Mapový klient SPOI.	60
8	Ukázka možnosti přístupu k datům OTM.	63
9	OTM jako dvouhvězdičková linked open data.	64
10	Datový model OTM.	65
11	Číselníky použité v OTM.	66
12	Open Land Use České republiky.	70
13	Open Land Use - as data.	70
14	Open Land Use - as a map.	71
15	OLU změny v modelu oproti INSPIRE.	72

Seznam zkratek

AOPK Agentura ochrany přírody a krajiny České republiky
API Application programming interface
BMIS Bezdrátový místní informační systém
CAD Computer-aided design
CC Classification of Types of Constructions
CC0 Creative Commons
CC BY Creative Commons - Uveďte původ
CC SA Creative Commons - Zachovejte licenci
CEGIS Center of Excellence for Geospatial Information Science
CKAN Comprehensive Knowledge Archive Network
CORINE Coordination of Information on the Environment
CPC Central Product Classification
CSV Comma-separated values
CSW Catalog Service for the Web
ČEZ České energetické závody (dříve)
ČHMÚ Český hydrometeorologický ústav
ČR Česká republika
ČSN Česká technická norma
ČSÚ Český statistický úřad
ČTÚ Český telekomunikační úřad
ČÚZK Český úřad zeměměřický a katastrální
DCAT Data Catalog Vocabulary
DGN “design”
DPZ Dálkový průzkum Země
DTM Digitální technická mapa
DXF Drawing Interchange Format
ECLAS European Commission Libraries Catalogue

ECMA European Computer Manufacturers Association
EEA European Environmental Agency
EPSG European Petroleum Survey Group
ESA European Space Agency, Evropská kosmická agentura
ESRI Environmental Systems Research Institute
ETRS European Terrestrial Reference System
EU European Union, Evropská unie
FAO Food and Agriculture Organization
FOAF Friend of a friend
G8 Group of Eight
GB Gigabyte
GEMET GEneral Multilingual Environmental Thesaurus
GEO Group on Earth Observation
GEOSS Global Earth Observation System of Systems
GIS Geografický informační systém
GML Geography Markup Language
GNU GPL GNU General Public License
HILUCS Hierarchical INSPIRE Land Use Classification System
HS-RS Help service - remote sensing
HTML HyperText Markup Language
IČ Identifikační číslo
IČO Identifikační číslo osoby
IETF Internet Engineering Task Force
INSPIRE INfrastructure for SPatial InfoRmation in Europe
IPR Institut plánování a rozvoje hlavního města Prahy
ISKN Informační systém katastru nemovitosti
ISO International Organization for Standardization, Mezinárodní organizace pro normalizaci
ISÚI Informační systém územní identifikace
JPG Joint Photographic Experts Group
JRC Joint Research Centre
JSON JavaScript Object Notation
JSON-LD JavaScript Object Notation for Linking Data
KML Keyhole Markup Language
LAEA Lambert azimuthal equal-area projection
LCC Lambert conformal conic projection
LČR Lesy České republiky
LOD Linked Open Data

LPIS Land Parcel Identification Systems
MARC MACHine-Readable Cataloging
METIS Metainformační systém
MHD Městská hromadná doprava
MOMC městský obvod/ městská část
MV ČR Ministerstvo vnitra České republiky
NAL National Agricultural Library
NGA Next Generation Access
NKOD Národní katalog otevřených dat
NUTS Nomenclature des Unites Territoriales Statistiques
OAI-PMH Open Archives Initiative Protocol for Metadata Harvesting
OD4D Open Data for Development
ODbL Open Database License
OGC Open Geospatial Consortium
ORP Obec s rozšířenou působností
OSN Organizace spojených národů
OTN Open Transport Net
OWL Web Ontology Language
OWS OGC Web Services
PDF Portable Document Format
PNG Portable Network Graphics
POI Point of Interest
POÚ Pověřený obecní úřad
PSI Public Sector Information
RCC Region Connection Calculus
RDF Resource Description Framework
RDFS Resource Description Framework Schema
RDMS Relational Database Management System
REST Representational state transfer
RGB red-green-blue; barevný model červená-zelená-modrá
ROB Registr obyvatel
ROS Registr osob
RPP Registr práv a povinností
RSS Rich Site Summary
RÚIAN Registr územní identifikace, adres a nemovitostí
RWE Rheinisch-Westfälische Elektrizitätswerke
ŘSD Ředitelství silnic a dálnic
S-JTSK Systém Jednotné trigonometrické sítě katastrální

SDI Spatial Data Infrastructure
SHP ESRI Shapefile
SKOS Simple Knowledge Organization System
SQL Structured Query Language
SPARQL SPARQL Protocol and RDF Query Language
SPOI Smart Points of Interest
SQL Structured Query Language
SWOT Strengths, Weaknesses, Opportunities, Threats
TIFF Tagged Image File Format
UMN University of Minnesota
UNBIS United Nations Bibliographic Information System
URI Uniform Resource Identifier
URL Uniform Resource Locator
USA United States of America, Spojené státy americké
USGS U.S. Geological Survey
UTM Universal Transverse Mercator
ÚHÚL Ústav pro hospodářskou úpravu lesů
UML Unified Modeling Language
ÚSES Územní systém ekologické stability
VF XML Výměnný formát XML
VFR Výměnný formát RÚIAN
VÚMOP Výzkumný ústav meliorací a ochrany půd. v.v.i.
VÚSC Vyšší územně-správní celek
W3C World Wide Web Consortium
WCS Web Coverage Service
WFS Web Feature Service
WGS World Geodetic System
WKT Well-known text
WMS Web Map Service, Webová mapová služba
XML Extensible Markup Language
ZSJ Základní sídelní jednotka
ZÚ Zeměměřický úřad
ZÚR Zásady územního rozvoje

Seznam příloh

Příloha A: Analytická část

Úvod

Metodika pro publikování prostorových informací ve formě otevřených dat¹ je koncipována jako sada postupů, které uživateli (subjektu publikujícímu prostorová data jako data otevřená) poskytne návod jak implementovat jednotlivé kroky (například volba licence, vazba na externí data, vhodný formát nebo tvorba identifikátorů) spojené s otevíráním dat, přičemž jsou zohledněna specifika prostorových dat (například velký rozsah dat, existence topologických vazeb nebo georeferencování).

Dokument byl vypracovaný kolektivem, který se skládal ze členů Katedry geomatiky² Fakulty aplikovaných věd Západočeské univerzity v Plzni a společnosti Help service remote sensing³. Jednotliví odborníci mají zkušenosti s otvíráním prostorových dat (například v rámci mezinárodních projektů SDI4Apps nebo SmartOpenData) a také se podíleli na tvorbě GeoInfoStrategie (jako klíčového dokumentu pro rozvoj geoinformatiky, včetně sad prostorových dat, v České republice) a vývoje i implementace datových specifikací a dalších součástí směrnice INSPIRE.

Pro tvorbu metodiky byly využity i články a další publikace členů autorského kolektivu - Čerba a Jedlička (2014), Čerba et al. (2016), Čerba a Jedlička (2016) - i texty další respektovaných expertů na otevřená data a jejich implementaci na doméně prostorových dat – Lebo a Williams (2010), Bauer a Kaltenböck (2011), Shadbolt et al. (2012), (*Open Data Consultancy. Final Report* 2013), Mráček a kolektiv (2014), Raamkumar, Thangavelu, a others (2015), Chlapek

¹V dalším textu bude kromě českého termínu, případně dalších českých synonym volná data nebo volně šiřitelná data, používáno i anglické (v českém prostředí etablované) spojení open data, přičemž oba výrazy budou chápány jako synonyma.

²<http://kgm.zcu.cz>

³<http://bnhelp.cz>

et al. (2015); Kavouras a Kokla (2007), Kovacs, Dolbear, a Goodwin (2007), Kritikos, Rousakis, a Kotzinos (2013), Kuhn, Kauppinen, a Janowicz (2014), Čepický et al. (2015).

Informace uvedené v tomto dokumentu jsou aktuální ke konci října 2016 (poslední revize byla provedena na konci listopadu 2016), což je zapotřebí brát v úvahu při implementaci metodiky, neboť oblast otevřených dat je velice dynamická. Dokument je zaměřený především na aktuální situaci v České republice, včetně současné legislativy, ale opírá se i o mezinárodní aktivity, mezi které patří především ISO normy, standardy organizací W3C nebo OGC, směrnice INSPIRE a příklady dobré praxe z různých mezinárodních projektů (podrobnosti jsou uvedeny v analytické části, která je v rámci metodiky publikována jako Příloha A).

Metodika se podrobně nezabývá otevřenými daty jako obecným fenoménem. Vlastnosti a výhody otevřených dat jsou detailně popsány v mnoha publikacích. V českém prostředí se jedná například o publikace Mráček a kolektiv (2014), Chlapek et al. (2015) nebo Čepický et al. (2015), z mezinárodních aktivit lze doporučit například Bauer a Kaltenböck (2011) nebo (*Open Data Consultancy. Final Report* 2013). Text metodiky akcentuje především specifika prostorových dat (viz níže uvedená struktura dokumentu). Jedná se především o technický dokument, který nemá ambici obhajovat koncept otevřených dat ani rozhodovat nebo doporučovat otevírání konkrétních datových sad (v tomto případě se jedná často o politické rozhodnutí spojené se změnou financování dané organizace, viz výstupy projektu TB0500MV002 Model financování datového fondu klíčových prostorových informací), ale ukazovat postupy, jak by bylo možné zvýšit otevřenost prostorových dat.

Metodika je určena především pro jednotlivé subjekty veřejné správy, které uvažují o otevřených prostorových datech z jakýchkoli pohnutek (to znamená i v případě, kdy si daná organizace plně neuvědomuje výhody otevřených dat, ale je povinna data v této formě publikovat). Navrhované postupy však mohou pro publikování nebo další využívání otevřených dat s výhodou využít i komerční nebo akademické organizace.

Práce je členěna tradičně. Kromě úvodní pasáže a závěrečného shrnutí obsahuje seznam používaných zkratk a jednoduchý slovníček základních pojmů a Přílohu A, která obsahuje analytickou část návrhu. Hlavní část metodiky je členěna podle tzv. pětihvězdičkového klasifikačního systému (Berners-Lee (2006)), který posuzuje otevřenost a provázanost dat podle pěti základních

parametrů (jde o kumulační systém, proto pro získání každého stupně musí být splněny všechny předchozí podmínky):

1. Data jsou publikována na webu pod otevřenou licenci.
2. Data jsou ve strojově čitelném formátu.
3. Data jsou v neproprietárním formátu.
4. Data jsou ukládána, dotazována a přenášena prostřednictvím standardů W3C (URI, HTTP, RDF a SPARQL)
5. Data jsou propojena s dalšími (externími) daty.

Jak bylo uvedeno výše, v jednotlivých kapitolách, které odpovídají jednotlivým hvězdičkám v klasifikačním systému, nebudou diskutovány obecné vlastnosti dat, ale především specifika a dílčí problémy spojené s prostorovými daty publikovanými v České republice (například zápis souřadnic, topologické vlastnosti nebo identifikátory pro registry veřejné správy). Každá kapitola obsahuje teoretickou část a na závěr sadu konkrétních doporučení pro prostorová data.

Konkrétní příklady aplikace open data přístupů jsou uvedeny v části Modelové aplikace, která obsahuje šest konkrétních a funkčních aplikací (nástrojů, datových sad), jež ilustrují jednotlivé aspekty prostorových otevřených dat:

- Metadata
- Ontologie VF XML - sémantická vrstva otevřených dat
- Smart Points of Interest - licence, využívání RDF pro prostorová data (body), tvorba slovníků
- Open Transport Map - propojení s INSPIRE registry, využívání RDF pro prostorová data (linie)
- Open Land Use Map - propojení s INSPIRE registry, využívání RDF pro prostorová data (plochy)
- Nepřímé geokódování (adresy) - formáty a slovníky

Kapitola 1

Slovník základních pojmů

Vzhledem k tomu, že metodika je zaměřena na prostorové informace a otevřená data, autoři považují za vhodné nejprve uvést definice těchto klíčových termínů, tak jak budou používány v analytické i návrhové části projektu.

Riziko používání nesprávných termínů je patrné například z publikace Jak otevírat data? (Mráček a kolektiv 2014), která se také zabývá otevíráním dat, včetně dat prostorových. Autoři, kteří jsou zřejmě nepoučení v oblasti prostorových dat, když používají naprosto nekorektní definici pro geodata - „Datasety mající reálnou kartografickou reprezentaci (je možné je zobrazit na mapě)“.

Prostorová informace, geoinformace, geografická informace (Šíma 2016)

1. informace týkající se jevů implicitně nebo explicitně přidružených k místu vztahenému k Zemi
2. znalost získaná jako výsledek syntézy, analýzy nebo integrace geografických dat¹

Prostorová data, geodata, geografická data (Šíma 2016)

1. data s implicitním nebo explicitním vztahem k místu na Zemi
2. počítačově zpracovatelná forma informace týkající se jevů přímo nebo nepřímo přidružených k místu na Zemi

¹Podle (ČSN P ISO/TS 19104 (979823) *Geografická informace - Terminologie* 2010) a Pražák a ostatní (2005–2016).

-
3. data identifikující geografickou polohu a charakteristiky přírodních a antropogenních jevů a hranic mezi nimi²

Otevřená data („298/2016 Sb. Zákon, kterým se mění některé zákony v souvislosti s přijetím zákona o službách vytvářejících důvěru pro elektronické transakce, zákon č. 106/1999 Sb., o svobodném přístupu k informacím, ve znění pozdějších předpisů, a zákon č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon), ve znění pozdějších předpisů" 2016)

Otevřenými daty se pro účely zákona 298/2016 Sb. rozumí informace zveřejňované způsobem umožňujícím dálkový přístup v otevřeném a strojově čitelném formátu, jejichž způsob ani účel následného využití není omezen a které jsou evidovány v národním katalogu otevřených dat.³

Otevřená data, volná, otevřená, otevřeně šiřitelná data (Šíma 2016)

data zveřejněná na internetu, která jsou úplná, snadno dostupná, strojově čitelná, používající standardy s volně dostupnou specifikací, zpřístupněná za jasně definovaných podmínek užití dat s minimem omezení a dostupná uživatelům při vynaložení minima možných nákladů⁴

Otevřená data; open government data (Open Knowledge Foundation)

- “Open” - freely used, reused and redistributed by material (data) is open if it can be anyone
- “Government data” produced or commissioned by government or data and information government controlled entities.

Otevřená data, open data (Hakrae 2012)

²Podle (ČSN P ISO/TS 19104 (979823) *Geografická informace - Terminologie* 2010), (Akční plán *Strategie rozvoje infrastruktury pro prostorové informace v České republice do roku 2020* 2015) a Pražák a ostatní (2005–2016).

³Poznámka autorů metodiky: V definici jsou nesmyslně kombinovány výrazy “data” a “informace” (otevřená data musí stále zůstat daty, tedy strojově čitelnými a formalizovanými entitami, a nemůže se jednat o informace). Navíc vazba na konkrétní katalogový produkt naprosto popírá princip otevřenosti.

⁴Omezení při poskytování dat (například požadavek na otevřený formát pro data i metadata) jsou popsána také v Předpisu č. 106/1999 Sb. Zákon o svobodném přístupu k informacím. Podle (Akční plán *Strategie rozvoje infrastruktury pro prostorové informace v České republice do roku 2020* 2015). Stejná definice je publikována i v sekci Otevřená data na stránce eGovernment MV ČR.

-
- Equal access for all
 - Licensing, legal issues
 - Transparency
 - Changing the way government works

Jak je patrné z předchozího textu, definice pojmu prostorová data jsou různorodé. V dalším textu bude pojmem otevřená data chápána definice z (*Akční plán Strategie rozvoje infrastruktury pro prostorové informace v České republice do roku 2020* 2015) ze dvou základních důvodů:

- Udržení kontinuity s GeoInfoStrategií a dalšími projekty zaměřenými na její implementaci.
- Vlastnosti otevřených dat, které jsou pro návrh metodiky klíčové, se ve všech uvedených definicích víceméně shodují.

Definice dalších pojmů jako jsou například metadata nebo licence budou uvedeny podle potřeby v této pasáži nebo dalším textu analytické nebo návrhové části.

Kapitola 2

Metodika (Návrhová část)

2.1 Licence a ochrana osobních údajů

2.1.1 Ochrana osobních údajů

Ochrana osobních údajů je v České republice regulována zákonem č. 101/2001 Sb., o ochraně osobních údajů a o změně některých zákonů, a dalšími právními předpisy. Zákon rozlišuje osoby, které zpracovávají osobní údaje jiných lidí (tzv. správce nebo zpracovatel osobních údajů) a dále osoby, jejichž osobní údaje správce a zpracovatelé zpracovávají (tzv. subjekty údajů). Správcům a zpracovatelům jsou při ochraně osobních údajů ukládány především povinnosti, zatímco subjektům údajů jsou dána práva.

Zákon č. 101/2001 Sb., o ochraně osobních údajů, se vztahuje na osobní údaje, které zpracovávají státní orgány, orgány územní samosprávy, jiné orgány veřejné moci, jakož i fyzické a právnické osoby (§3 odst. 1), a vztahuje se na veškeré zpracovávání osobních údajů, ať k němu dochází automatizovaně nebo jinými prostředky (§3 odst. 2). Tento zákon se nevztahuje na zpracování osobních údajů, které provádí fyzická osoba výlučně pro osobní potřebu (§3 odst. 3), a nevztahuje se na nahodilé shromažďování osobních údajů, pokud tyto údaje nejsou dále zpracovávány (§3 odst. 4.).

Osobní údaj je v § 4 písm. a) zákona č. 101/2001 Sb., vymezen jako jakákoliv informace týkající se určeného nebo určitelného subjektu údajů. Subjekt

údajů se považuje za určený nebo určitelný, jestliže lze subjekt údajů přímo či nepřímo identifikovat zejména na základě čísla, kódu nebo jednoho či více prvků, specifických pro jeho fyzickou, fyziologickou, psychickou, ekonomickou, kulturní nebo sociální identitu. Citlivým údajem je osobní údaj vypovídající o národnostním, rasovém nebo etnickém původu, politických postojích, členství v odborových organizacích, náboženství a filozofickém přesvědčení, odsouzení za trestný čin, zdravotním stavu a sexuálním životě subjektu údajů a jakýkoliv biometrický nebo genetický údaj subjektu údajů (§ 4 písm. b) zákon č. 101/2001 Sb.). Anonymním údajem je takový údaj, který buď v původním tvaru, nebo po provedeném zpracování nelze vztáhnout k určenému nebo určitelnému subjektu údajů (§ 4 písm. c) zákona č. 101/2001 Sb.).

Zákon č. 101/2001 Sb., o ochraně osobních údajů, dále reguluje zpracování osobních údajů, čímž jsou jakákoliv operace nebo soustava operací, které správce nebo zpracovatel systematicky provádějí s osobními údaji, a to automatizovaně nebo jinými prostředky. Zpracováním osobních údajů se rozumí zejména shromažďování, ukládání na nosiče informací, zpřístupňování, úprava nebo pozměňování, vyhledávání, používání, předávání, šíření, zveřejňování, uchovávání, výměna, třídění nebo kombinování, blokování a likvidace (§ 4 písm. e). Zákon definuje shromažďování osobních údajů jako systematický postup nebo soubor postupů, jehož cílem je získání osobních údajů za účelem jejich dalšího uložení na nosič informací pro jejich okamžité nebo pozdější zpracování (§ 4 písm. f), a uchovávání osobních údajů je udržování údajů v takové podobě, která je umožňuje dále zpracovávat (§ 4 písm. g).

Systematicky prováděnou činností je zejména tvorba různých databází a rejstříků, které umožňují s osobními údaji dále pracovat. V praxi v podstatě jakékoli automatizované zpracování, které například umožňuje vyhledávání jednotlivých údajů, bude zpracováním osobních údajů.

2.1.2 Otevírání dat ve vztahu k zákonu o ochraně utajovaných skutečností.

Ochrana státního tajemství je zásadní součástí bezpečnostní politiky státu, a v praxi se s touto problematikou setkává řada úřadů a organizací. Ačkoli zdaleka nejde o problematiku novou, přístup k ní musí být v době budování informační společnosti zcela nový. Ochrana utajovaných skutečností nemůže být založena

pouze na prostém výčtu činností, které fyzická nebo právnická osoba musí udělat, aby naplnila podmínky pro získání určitého osvědčení či potvrzení. Zásadní důležitost a význam má především bezpečnost podniků, úřadů a informací (tedy nikoli pouze utajovaných skutečností) v dnešním světě, zejména například ve sféře obchodu, obrany a bezpečnostní problematiky. Úspěšné firmy již dnes pochopily, že vyřešením bezpečnostních otázek přispívají nejenom k úspěšnému rozvoji a stabilitě konkrétní organizace nebo instituce, ale také k větší stabilitě celé společnosti i k rozvoji celkové bezpečnostní kultury.

Vytvořením Schengenského prostoru byly právní předpisy tvořící základ Schengenské spolupráce začleněny do institucionálního rámce Evropské unie (EU), čímž byla Schengenská spolupráce přesunuta na úroveň Evropské unie. Tím přijímané předpisy EU představují unijní právní instrumenty platné pro všechny členské státy unie. Na úrovni ČR v souladu s právem Evropské unie upravuje práva a povinnosti v oblasti ochrany utajovaných skutečností jednak zákon č. 480/2004 Sb., o některých službách informační společnosti a o změně některých zákonů (zákon o některých službách informační společnosti) a Úmluva o ochraně osob se zřetelem na automatizované zpracování osobních dat č. 108, vyhlášená pod č. 115/2001 Sb. m. s.

Problematiku zpracování osobních údajů a podmínky pro předávání osobních údajů do jiných států a jejich využití ošetřuje Zákon č. 101/2000 Sb. Zákon o ochraně osobních údajů a o změně některých zákonů. Tento zákon se vztahuje na osobní údaje, které zpracovávají státní orgány, orgány územní samosprávy, jiné orgány veřejné moci, jakož i fyzické a právnické osoby. Zákon se vztahuje na veškeré zpracovávání osobních údajů, ať k němu dochází automatizovaně nebo jinými prostředky, nevztahuje se na zpracování osobních údajů, které provádí fyzická osoba výlučně pro osobní potřebu a na nahodilé shromažďování osobních údajů, pokud tyto údaje nejsou dále zpracovávány. Výjimky z působení zákona (bezpečnost ČR, obrana ČR, veřejný pořádek a vnitřní bezpečnost, odhalování trestné činnosti a stíhání trestných činů, výkon kontroly, dozoru, dohledu a regulace spojené s výkonem veřejné moci) ošetřuje §3, odst. 6 zákona č. 101/2000 Sb. Zákon definuje shromažďováním osobních údajů jako systematický postup nebo soubor postupů, jehož cílem je získání osobních údajů za účelem jejich dalšího uložení na nosič informací pro jejich okamžité nebo pozdější zpracování. Zpracováním osobních údajů se rozumí jakákoliv operace nebo soustava operací, které správce nebo zpracovatel systematicky provádějí s osobními údaji, a to automatizovaně nebo jinými prostředky. Zpracováním osobních údajů se rozumí zejména shromažďování,

ukládání na nosiče informací, zpřístupňování, úprava nebo pozměňování, vyhledávání, používání, předávání, šíření, zveřejňování, uchovávání, výměna, třídění nebo kombinování, blokování a likvidace. Uchovávání osobních údajů je definováno jako udržování údajů v takové podobě, která je umožňuje dále zpracovávat. Zveřejněným osobním údajem se v tomto zákoně rozumí osobní údaj zpřístupněný zejména hromadnými sdělovacími prostředky, jiným veřejným sdělením nebo jako součást veřejného seznamu.

Pro oblast otevírání dat je zcela zásadní Zákon č. 412/2005 Sb. Zákon o ochraně utajovaných informací a o bezpečnostní způsobilosti. Tento zákon upravuje zásady pro stanovení informací jako informací utajovaných, podmínky pro přístup k nim a další požadavky na jejich ochranu, zásady pro stanovení citlivých činností a podmínky pro jejich výkon a s tím spojený výkon státní správy.

2.1.3 Práva duševního vlastnictví a otevřené licence

Problematika publikování otevřených dat je úzce spojena s jejich ochranou z pohledu duševního vlastnictví. Pro publikování otevřených dat je potřeba nastavit podmínky takové, aby data bylo možné snadno sdílet, upravovat a dále šířit a to i za hranicemi státu. Z mezinárodního pohledu můžeme práva duševního vlastnictví rozdělit na:

1. Autorská práva (Copyright)
2. Databázová práva
3. Práva průmyslového vlastnictví (např. patenty, obchodní známky a průmyslové vzory).

V legislativě České republiky jsou první dvě kategorie zakotveny v zákoně č. 121/2000 Sb., o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon). Práva průmyslového vlastnictví jsou zakotvena ve více dílčích právních předpisech.

Z hlediska publikování otevřených informací ve formě otevřených dat jsou to právě autorská a databázová práva, které je nutné analyzovat před samotnou publikací dat. Podmínky užití dat jsou dány tzv. licencí. Obecná ustanovení k problematice licencí jsou regulována zákonem č. 89/2012 Sb., občanský zákoník v §§ 2358 až 2370. Zvláštní ustanovení pro licence k předmětům chráněným autorským zákonem jsou uvedena v §§ 2371 až 2377.

Licence upravuje, jak lze s daty nakládat. Může omezit např. způsob využití dat či časový limit, kdy lze data využít. Licence může být ujednána formou licenční smlouvy mezi vlastníkem dat a jejich uživatelem. Licence může být ale také přímo spojena s datovou sadou a každý uživatel musí podmínky dané licencí dodržet. Vlastník dat může přidělit různé licence (podmínky užití) pro různé skupiny uživatelů (např. pro nekomerční využití je možné využít data bez poplatku, pro komerční účely jsou data zpoplatněna).

V mezinárodním kontextu se užívá pojem otevřená licence (open licence), který je v souladu s tzv. Otevřenou definicí¹:

„Znalosti jsou otevřené, pokud k nim může kdokoliv volně přistupovat, používat je, upravovat a sdílet je – za podmínek, které nejvíce zabezpečí zachování původu, tedy vlastnictví, a otevřenosti.“

Otevřené licence musí umožnit uživatelům nejen přístup k datům, ale i možnost je znovu využít, upravovat, kombinovat s jinými daty či dále šířit, a to pro komerční i nekomerční účely. Otevřená definice umožňuje nastavit omezení uvedení původu (Attribution) a zachování licence (Share Alike).

Sdělení Evropské komise 2014/C 240/01 doporučuje používat normované mezinárodní otevřené licence Creative Commons² namísto vytváření vlastních licencí. Komise doporučuje použít CC0 Public Domain³, která „zajišťuje uživatelům plnou flexibilitu a omezuje komplikace spojené s prací s četnými licencemi, které mohou obsahovat protichůdná ustanovení.“ Creative Commons nabízí i otevřené licence s omezeními CC-BY (attribution) a CC-BY-SA (attribution and share-alike).

Obecný trend ve zpřístupňování informací veřejné sféry jako otevřená data je nastavení podmínek užití s co nejméně omezeními, ideálně bez jakýchkoliv omezení, tj. CC0 Public Domain. To je i obecné doporučení této metodiky.

Další informace o podmínkách užití datových sad můžete nalézt na stránkách <http://opendata.gov.cz/cinnost:stanoveni-podminek-uziti> zpracované obecně pro otevřená data.

¹<http://opendefinition.org/od/2.0/cz/>

²<https://creativecommons.org/>

³<https://creativecommons.org/publicdomain/zero/1.0/deed.cs>

2.2 Metadata

Při publikování otevřených dat mají metadata klíčovou roli.

- Představují dokumentaci dat, která je nutná pro jejich správnou interpretaci.
- Pomocí metadat jsou vytvářeny katalogy, ve kterých je možné data podle určitých parametrů vyhledávat.

Z těchto důvodů by každá otevřená data měla být vybavena metadaty.

2.2.1 Shrnutí požadavků na metadata

1. Vzhledem k tomu, že řada projektů v oblasti geoinformatiky se řídí směrnicemi INSPIRE, doporučuje⁴ se primárně vytvářet všechna metadata prostorových dat v souladu s nařízeními INSPIRE (INSMD INSPIRE 2008, INSNS INSPIRE (2009), I. M. I. INSPIRE (2016), INSPIRE (2010)) a národním metadatovým profilem (Metadata 2015).
2. Pro uplatnění přístupů sémantického webu se doporučuje vytvářet metadata v souladu s posledními technickými návody pro metadata (I. M. I. INSPIRE 2016) s maximálním použitím URI pro klíčová slova, souřadnicové referenční systémy a další metadatové prvky.
3. Pro použití v katalogích otevřených dat by měla být metadata souběžně publikována ve formátu GeoDCAT-AP (Bechhofer a Miles 2009) v souladu s evropskými nařízeními. Je možnost využít k automatické konverzi např. nástroje na Národním geoportálu INSPIRE⁵, aby metadata nemusela být pořizována souběžně ve dvou formátech.
4. Aby bylo umožněno vyhledání, je třeba metadata publikovat v katalogu. Primárně se doporučuje Národní katalog INSPIRE⁶, sekundárně Národní katalog otevřených dat⁷. Technicky bude zajištěna automatická

⁴Doporučení v tomto dokumentu nepředstavují obecně platná pravidla nebo regule (například legislativní či normativní doporučení), ale výsledek analytické a syntetické práce autorů dokumentu.

⁵<https://geoportal.gov.cz/web/guest/metadata/geodcat/>

⁶<http://geoportal.gov.cz>

⁷<http://portal.gov.cz/portal/rejstriky/data/97898/>

synchronizace⁸ mezi těmito katalogy. Je vhodné se informovat u jejich správců.

5. Pro správnou interpretaci dat je nutno poskytovat dokumentaci jejich struktury (popis atributů, jejich význam a domény) alespoň ve formě textového dokumentu.

2.2.2 Detailní popis

V posledních technických návodech pro metadata INSPIRE (I. M. I. INSPIRE 2016) je vidět výrazný posun k sémantickému webu. Řadu položek je možnost prezentovat ve formě URI, které jsou v ISO 19139 implementovány pomocí elementu `gmx:Anchor`. I když je umožněno nadále stále uvádět některé položky pomocí řetězců, doporučuje se upřednostnit tento přístup pro budoucí kompatibilitu.

Zde předkládané technické řešení vychází ze standardu GeoDCAT a nástrojů poskytovaných JRC pro převod z metadat ISO 19139 (INSPIRE) do GeoDCAT, ovšem rozšiřuje tento přístup o následující aspekty:

- Podpora multilinguálního řešení
- Rozšíření konceptu o další požadavky definované tímto projektem, např. adresa apod.
- Rozšíření mapování vazeb mezi metadatovými záznamy
- Podpora zpětného mapování klíčových slov INSPIRE a citací dokumentů INSPIRE v textovém formátu na URI

2.2.2.1 Klíčová slova

Doporučuje se kódovat témata INSPIRE pomocí URI ve tvaru uváděném v INSPIRE registry.

Příklad v ISO 19139:

```
<gmd:keyword>  
  <gmx:Anchor xlink:href=  
    "http://inspire.ec.europa.eu/theme/lc">
```

⁸Orientačně od dubna 2017.

```
Krajinný pokryv</gmx:Anchor>
</gmd:keyword>
```

Příklad v GeoDCAT:

```
<dcat:theme rdf:resource="http://inspire.ec.europa.eu/theme/lc"/>
```

Doporučuje se podobně kódovat další klíčová slova, pokud pocházejí z nějakého tezauru podle SKOS (Bechhofer a Miles 2009), např. GEMET

Příklad v ISO 19139:

```
<gmd:keyword>
  <gmx:Anchor xlink:href="http://www.eionet.europa.eu/gemet/concept/
4648">Krajina</gmx:Anchor>
</gmd:keyword>
```

Příklad v GeoDCAT:

```
<dcat:theme rdf:resource="http://www.eionet.europa.eu/gemet/
concept/4648"/>
```

2.2.2.2 Aplikační schéma / struktura dat

Pro účely publikace otevřených dat je tato informace zásadní ke správné interpretaci dat, proto musí být pro uživatele lehce přístupná⁹.

V ISO 19139 je možno strukturu dat dokumentovat buď

- odkazem na Feature katalog (v ISO 19115-1:2014 je možné Feature katalog přímo vložit do metadat)
- V `MD_ApplicationSchemaInformation`, kde je odkaz na UML model nebo jinou dokumentaci

Doporučuje se poskytovat dokumentaci dat ve formě odkazu na dokument s detailním popisem struktury dat. V obou případech se doporučuje kódovat tento odkaz v GeoDCAT pomocí `dcat:documentation`.

V budoucnu by struktura mohla být dokumentována pomocí ontologického popisu.

⁹INSPIRE dokumentaci struktury dat v metadatech neřeší. Struktura je dokumentována v popisu jednotlivých datových témat.

Příklad v ISO 19139: Aplikační schéma

```
<gmd:MD_ApplicationSchemaInformation>
  <gmd:name>
    <gmd:CI_Citation>
      <gmd:title>
        <gco:CharacterString>Textová dokumentace v PDF
      </gco:CharacterString>
      </gmd:title>
      <gmd:date>
        <gmd:CI_Date>
          <gmd:date>
            <gco:Date>2016</gco:Date>
          </gmd:date>
          <gmd:dateType>
            <gmd:CI_DateTypeCode codeList="..."
codeListValue="creation"/>
          </gmd:dateType>
        </gmd:CI_Date>
      </gmd:date>
    </gmd:CI_Citation>
  </gmd:name>
  <gmd:schemaLanguage gco:nilReason="inapplicable" />
  <gmd:constraintLanguage gco:nilReason="inapplicable" />
  <gmd:graphicsFile>
    <gco:Binary src="http://myfile.cz"/>
  </gmd:graphicsFile>
</gmd:MD_ApplicationSchemaInformation>
```

Příklad v GeoDCAT:

```
<dcat:documentation>
  <foaf:page rdf:resource="http://myfile.cz"/>
</dcat:documentation>
```

2.2.2.3 Licence

Licence je důležitým údajem pro nakládání s daty. Pro geodata se doporučuje zejména využití licencí Creative Commons, pokud bude nutné též licencí INSPIRE.

Příklad kódování licencí Creative Commons bude vypadat následovně:

Příklad v ISO 19139:

```
<gmd:useLimitation>
  <gmx:Anchor xlink:href=
"http://creativecommons.org/licenses/by-sa/4.0/">
CC BY-SA 4.0</gmx:Anchor>
</gmd:useLimitation>
```

Příklad v GeoDCAT:

```
<dcat:Distribution>
  <dct:license rdf:resource=
"http://creativecommons.org/licenses/by-sa/4.0/">
</dcat:Distribution>
```

Odkaz na licenci INSPIRE by měl mířit na národní geoportál INSPIRE, např. : https://geoportal.gov.cz/c/document_library/get_file?uuid=e8e493a6-f787-439c-9409-757cbeb05ae5&groupId=10138.

Příklad v ISO 19139:

```
<gmd:useLimitation>
  <gmx:Anchor xlink:href="https://geoportal.gov.cz/
c/document_library/get_file?uuid=e8e493a6-f787-439c-
9409-757cbeb05ae5&groupId=10138">
Základní licence INSPIRE</gmx:Anchor>
</gmd:useLimitation>
```

Příklad v GeoDCAT:

```
<dcat:Distribution>
  <dct:licence rdf:resource=
"https://geoportal.gov.cz/c/document_library/
get_file?uuid=e8e493a6-f787-439c-9409-
```

```
757cbeb05ae5&groupId=10138"/>
</dcat:Distribution>
```

2.2.2.4 Adresy

Doporučuje se, aby adresy byly uvedeny nejen v textovém tvaru, ale též ve formě URI. Problémem je, že v ČR neexistuje dosud systém URI pro adresy. Veřejná správa by tento systém měla vytvořit a kodifikovat. Více v kapitole o identifikátorech.

Doporučuje se uvádět URI adresy v položce metadat **deliveryPoint** a souběžně vyplnit další položky adresy v textovém tvaru na odpovídající místa normy (viz příklad).

Příklad v ISO 19139:

```
<gmd:CI_Address>
  <gmd:deliveryPoint>
    <gmx:Anchor xlink:href="http://data.gov.cz/
reg/RUIAN/ADR/123456"> Pod sídlištěm 9 </gmx:Anchor>
  </gmd:deliveryPoint>
  <gmd:city><gco:CharacterString>Praha
</gco:CharacterString></gmd:city>
  <gmd:postalCode>
    <gco:CharacterString>182 11 Praha 8
</gco:CharacterString>
  </gmd:postalCode>
  <gmd:country>
    <gco:CharacterString>Česká republika
</gco:CharacterString>
  </gmd:country>
  <gmd:electronicMailAddress>
    <gco:CharacterString>cuzk.helpdesk@cuzk.cz
</gco:CharacterString>
  </gmd:electronicMailAddress>
</gmd:CI_Address>
```

Příklad v GeoDCAT:

```
<vcard:hasAddress rdf:resource="http://data.gov.cz/
reg/RUIAN/ADR/123456"/>
```

Poznámka: Patrně by bylo vhodné duplicitně uvádět adresu i textově i na tomto místě.

2.2.2.5 Organizace

Podobně jakou u adres se doporučuje uvádět URI organizace v elementu `organisationName` založenou na budoucích základních registrech (více v kapitole o identifikátorech).

Příklad v ISO 19139:

```
<gmd:CI_ResponsibleParty>
  <gmd:organisationName>
    <gmx:Anchor xlink:href="
http://data.gov.cz/reg/ROS/123456">Český úřad ...
  </gmx:Anchor>
</gmd:organisationName>
</gmd:CI_ResponsibleParty>
```

V GeoDCAT:

```
<dcat:contactPoint rdf:resource=
"http://data.gov.cz/reg/ROS/123456"/>
```

2.2.2.6 Prostorový rozsah dat

Směrnice INSPIRE vyžaduje pouze ohraničující obdélník v podobě geografických souřadnic. Doporučuje se přidat ještě element obsahující URI odpovídající prostorové jednotky ze základních registrů¹⁰ (alternativně je možné použít např. GeoNames.org). Pokud je rozsah odpovídající území České republiky, bude použita tato URI (v souladu s DCAT): `http://publications.europa.eu/resource/authority/country/CZE`.

Příklad v ISO 19139:

¹⁰Až budou stanoveny.

```

<gmd:EX_Extent>
  <gmd:geographicElement>
    <gmd:EX_GeographicDescription>
      <gmd:geographicIdentifier>
        <gmd:MD_Identifier>
          <gmd:code>
            <gco:CharacterString>
              http://data.gov.cz/reg/RUIAN/OBCE/123456
            </gco:CharacterString>
          </gmd:code>
        </gmd:MD_Identifier>
      </gmd:geographicIdentifier>
    </gmd:EX_GeographicDescription>
  </gmd:geographicElement>
</gmd:EX_Extent>

```

Příklad v GeoDCAT:

```

<dct:spatial rdf:resource=
"http://data.gov.cz/reg/RUIAN/OBCE/123456"/>

```

2.2.2.7 Účel vytvoření

Účel vytvoření je významnou položkou pro zjišťování, z jakého titulu (důvodu) jsou určitá data pořizována ve veřejné správě. Proto je textové vyjádření nedostatečné a bylo by vhodné použít URI na odpovídající zákon či jiný předpis, pokud by byla tato data k dispozici jako LOD. Položka není v ISO 19139 násobná, proto je k diskusi zda je vhodné adresovat příslušnou legislativu přímo nebo použít odkaz na nějaký registr, odkud budou teprve další odkazy na jednotlivé zákony.

Možné způsoby mapování do DCAT (k diskusi):

```

dct:conformsTo
dcat:theme
dct:relation

```

Příklad v ISO 19139:

```

<gmd:purpose>

```

```
<gmx:Anchor xlink:href=
"http://data.gov.cz/reg/LEG/123456">
Zde citace zákona.</gmx:Anchor>
</gmd:purpose>
```

Příklad v GeoDCAT:

```
<dct:conformsTo rdf:resource=
"http://data.gov.cz/reg/LEG/123456"/>
```

2.2.2.8 Modelování vazeb

Vazby mezi daty navzájem a daty a službami jsou informačně významné pro interpretování metadat. Zvláště v prostředí LOD je modelování těchto vazeb důležité. Doporučuje se zohlednit tyto vazby (v popisu se vychází z aktuálního metadatového popisu):

1. Nadřízená datová sada (např. pokud je datová sada součástí série - katastrální mapy apod.)
2. Podřízené datové sady (symetrické k 1. V tomto případě ale může existovat i několik tisíc datových sad - jak u ČÚZK - je tedy třeba uvážit velikost metadatového záznamu)
3. Které datové sady jsou použité službou
4. Které služby používají datovou sadu (symetrické k 3.)

Pozn.: Technicky není možné zajistit modelování všech vazeb při převodu jednoho metadatového záznamu externím programem, které vyžaduje implementaci přímo v metadatovém katalogu, protože např. parentIdentifier poskytuje pouze identifikátor záznamu a neposkytuje odkaz na nařízený metadatový záznam (více u jednotlivých případů).

ad 1.

Je definován v GeoDCAT - ovšem zatím ve fázi diskuse. Popis zahrnuje jak odkaz na metadatový záznam ve tvaru GeoDCAT tak vlastní identifikátor dat. Není možno jej generovat externím konvertorem.

Příklad v ISO 19139:

```
<gmd:parentIdentifier>
<gco:CharacterString>id-rodíčovských-metadat
```

```
</gco:CharacterString>
</gmd:parentIdentifier>
```

Pozn.: ISO 19139 neposkytuje odkaz na nadřazený metadatový záznam, ten je k dispozici až v ISO 19115-3:2014)

Příklad v DCAT:

```
<dct:isPartOf rdf:parseType="Resource">
<dct:identifier rdf:datatype=
"http://www.w3.org/2001/XMLSchema#string">
dataset-id</dct:identifier>
<foaf:isPrimaryTopicOf>
<dcat:CatalogRecord rdf:about=
"url-rodicovských-metadat-DCAT"/>
</foaf:isPrimaryTopicOf>
</dct:isPartOf>
```

Pozn.: `dct:identifier` je generován z `identificationInfo/*/citation/*/identifier` ad 2.

Není uváděn v GeoDCAT. Mapuje se podobně jako 1, ovšem pomocí `dct:hasPart`. Není možno generovat externím konvertorem. V ISO není tato vazba modelována, je však možno ji nalézt v katalogu dotazem zpětně pomocí `parentIdentifier`.

Příklad v DCAT:

```
<dct:hasPart rdf:parseType="Resource">
<dct:identifier rdf:datatype=
"http://www.w3.org/2001/XMLSchema#string">
dataset-id</dct:identifier>
<foaf:isPrimaryTopicOf>
<dcat:CatalogRecord rdf:about=
"url-dceřiných-metadat-DCAT"/>
</foaf:isPrimaryTopicOf>
</dct:hasPart>
```

Pozn: V případě většího množství podřazených záznamů (X000 - X0000), jak je tomu v případě ČÚZK, by mohlo dojít k extrémnímu nárůstu velikosti souboru a případném zahlcení programu.

ad 3.

Podle metodiky v GeoDCAT. Tam je modelováno pomocí **hasPart**, což se nepovažuje za přesné. Lze generovat přímo i v externím konvertoru vzhledem k přítomnosti URL ve zdrojových metadatech.

Příklad v ISO 19139:

```
<srv:operatesOn xmlns:xlink=
"http://www.w3.org/1999/xlink"
xlink:type="simple"
xlink:href="url-metadat-dat-ISO"/>
```

Příklad v DCAT:

```
<dct:hasPart rdf:parseType="Resource">
<dct:identifier rdf:datatype=
"http://www.w3.org/2001/XMLSchema#string">
id-dat</dct:identifier>
<foaf:isPrimaryTopicOf>
<dcat:CatalogRecord rdf:about="url-metadat-dat-DCAT"/>
</foaf:isPrimaryTopicOf>
</dct:hasPart>
```

ad 4.

Není definován v GeoDCAT, V ISO není přímo definován, ale je možné jej modelovat dotazem do katalogu. Je otázkou, zda používat **dct:isPartOf**. Identifikátor služby není v INSPIRE požadován, proto zde nemusí být uveden

Příklad v DCAT:

```
<dct:isPartOf rdf:parseType="Resource">
<dct:identifier rdf:datatype=
"http://www.w3.org/2001/XMLSchema#string">
id-sluzby</dct:identifier>
<foaf:isPrimaryTopicOf>
<dcat:CatalogRecord rdf:about=
"url-metadat-sluzby-DCAT"/>
</foaf:isPrimaryTopicOf>
</dct:isPartOf>
```

2.2.3 Příklady

1. Vzorový soubor metadat INSPIRE¹¹
2. Odpovídající vzorový soubor metadat ve formátu GeoDCAT-AP¹²

2.2.4 Postup při publikaci otevřených dat

2.2.4.1 Scénář 1 - Publikování přes Národní katalog INSPIRE

1. Umístěte data na veřejně přístupnou adresu nebo vytvořte veřejně přístupnou webovou službu.
2. Vytvořte odpovídající metadata ve formátu INSPIRE (ISO 19139) buď ve Vašem vlastním metadatovém editoru nebo použijte editor na [Národním geoportálu INSPIRE¹³.
3. Publikujte metadata na národním geoportálu INSPIRE¹⁴.
 - a) Buď přímo na portálu nebo importem souboru do metadatového editoru.
 - b) Harvestingem z Vašeho připojeného katalogu .
4. Z Národního geoportálu INSPIRE by měly být Vaše záznamy do 24 hodin replikovány na Národním katalogu otevřených dat. Pokud se tak nestane, prosím, kontaktujte technickou podporu.

2.2.4.2 Scénář 2 - Publikování přes Národní katalog otevřených dat

V případě, že z nějakého důvodu nemůžete publikovat Vaše metadata na Národním geoportálu INSPIRE postupujte takto:

1. Publikujte data jako ve scénáři 1.
2. Vytvořte metadata jako ve scénáři 1.

¹¹<http://geoportal.gov.cz/res/example-anchor.xml>

¹²<http://geoportal.gov.cz/res/example-dcat.xml>

¹³<https://geoportal.gov.cz/web/guest/metadata/create/>

¹⁴Pokud nejste registrováni na Národním geoportálu INSPIRE, je možné se registrovat zde: <https://geoportal.gov.cz/web/guest/create-account> Postupujte podle instrukcí zde uvedených.

-
3. Metadatový záznam v INSPIRE formátu můžete převést pomocí ON-LINE nástroje¹⁵ do formátu GeoDCAT.
 4. Poté publikujte metadata ve formátu GeoDCAT na Národní katalog otevřených dat¹⁶.

2.3 Formáty prostorových dat

Formáty prostorových dat lze rozdělit dle kritérií pětihvězdičkového schématu pro otevřená data na formáty otevřené a proprietární. Za otevřené formáty lze považovat takové formáty, které mají dostupnou specifikaci, jež je publikována nezávislou standardizační organizací (např OGC nebo ISO). Často používané formáty, včetně ukázky softwaru na jejich zpracování, jsou znázorněny v Tabulce 1.

2.3.1 Přehled používaných formátů a software

Proprietárních formátů, které jsou však v komunitě GIS hojně používány, je celá řada. Ačkoliv nesplňují požadavky pětihvězdičkového schématu, jsou mnohdy hojně využívány. Nevýhodou je možná změna jejich specifikace, případně uzavření a znepřístupnění této specifikace. Nejpožívanější formáty včetně příslušné organizace jsou znázorněny v Tabulce 1.

Tabulka 1: Příklady formátů v GIS.

Název formátu	Přípona	Příklad software
Arc/Info & Space Imaging BIP	.bip	ArcMap
Arc/Info & Space Imaging BSQ	.bsq	ArcMap
AutoCad Drawing (DWG)	.dwg	AutoCad
AutoCAD DXF	.dxf	AutoCad, ogr
MicroStation Design File (DGN)	.dgn	Microstation, ogr
ER Mapper	.ers	ArcMap

¹⁵<https://geoportal.gov.cz/web/guest/metadata/geodcat>

¹⁶<http://portal.gov.cz/portal/rejstriky/data/>

Název formátu	Přípona	Příklad software
ERDAS 7.5 GIS	.gis	ERDAS, ArcMap
ERDAS 7.5 LAN	.lan	ERDAS, ArcMap
ESRI binary interleaved format (BIL)	.bil	ArcMap
Floating-Point Grid (FLT)	.flt	ArcMap
Golden Software Grid	.grd	ArcMap
GRIB	.grb	ArcMap
GSat	.gff	ArcMap
GXF (Grid eXchange File)	.gxf	ArcMap
IDRISI Raster	.rst	ArcMap
Intergraph CIT	.cit	Geomedia, ArcMap
Intergraph COT	.cot	Geomedia, ArcMap
Intergraph CTB	.ctb	Geomedia, ArcMap
Intergraph CTC	.ctc	Geomedia, ArcMap
Pictometry Oblique Image (PSI)	.psi	ArcMap
SRTM HGT	.hgt	ArcMap
Terragen Heightfield (TER)	.ter	ArcMap
USGS ASCII DEM	.dem	ArcMap
USGS SDTS DEM	.ddf	ArcMap
ER Mapper ECW	.ecw	ArcMap, gdal
Erdas Imagine IMG	.img	ArcMap, gdal
File Geodatabase (FileGDB)	.gdb	ogr
Ascii XYZ	.xyz	gdal
MrSID	.sid	gdal
Global Navigation Chart (GNC)	.gn1, .gn2, .gn3, ...	gdal, ArcMap

Název formátu	Přípona	Příklad software
Naval Range Operating Area Chart	.oa1, .oa2, .oa3, ...	gdal, ArcMap
PCI Geomatics Database File	.pix	gdal, ArcMap
Raster Product (RPF)	.rpf	gdal, ArcMap
National Imagery Transmission Format (NITF)	.ntf, .nitf	gdal, ArcMap
Flexible Image Transport System (FITS)	.fits, .fit, .fts	gdal
GPS eXchange Format (GPX)	.gpx	gpx, ogr
Compressed Lidar File (LIDAR)	.laz	lastools
Lidar (LIDAR)	.las	lastools
Atlas BNA	.bna	ogr
GeoConcept Text Export (GXT)	.gxt	ogr
GMT Ascii Vector	.gmt	ogr

2.3.2 Hlavní otevřené datové formáty pro prostorová data

Mezi nejdůležitější otevřené formáty pro vektorová data používané v oblasti otevřených dat patří především KML, GML, SHP a GeoJSON. Formáty KML, GML a GeoJSON jsou specifikované konsorciem OGC. Formát SHP je specifikován firmou ESRI jako otevřený a v oblasti GIS je obecně zavedeným a používaným formátem.

2.3.2.1 Comma Separated Values (CSV)

Bezpochyby nejjednodušší způsob poskytnutí dat je textový dokument. Tzv CSV je obvykle dokument, kde v první řádce jsou uvedeny názvy jednotlivých sloupců a následují řádky s jednotlivými datovými záznamy odděleny čárkou případně jiným způsobem. Důležitou součástí dat je popis jednotlivých atributů a informace o souřadnicové systému prostorových souřadnic.

```
x,y,accident_severity,speed_limit,timestamp,vehicle_type
126.6870848,83.96586725,2,30,1451591400,7
126.63756871,83.99031819,3,30,1451576280,6
126.66713458,83.95373734,3,30,1451563260,8
126.66357049,84.01691118,3,30,1451544540,9
126.60777102,84.01112928,3,30,1451502000,9
126.60777102,84.01112928,3,30,1451502000,9
126.69904142,84.01444347,3,30,1451501220,9
```

2.3.2.2 Keyhole Markup Language - KML/KMZ

KML je formát postavený na XML (KMZ je komprimovaná varianta v zip formátu). Původně byl formát spjat s aplikací Google Earth a mapovými produkty firmy Google. V dnešní době je formát standardem OGC a jeho plný popis a aktuální verzi lze nalézt na stránkách OGC.

Formát je vhodný pro jednoduchá data, především z důvodu podpory v GIS aplikacích i aplikacích pro širokou veřejnost. Na druhou stranu KML není vhodný pro uchování komplexních dat, kde se předpokládá jejich využití v GIS analýzách.

Ukázka KML:

```
<Placemark>
  <name>Plzeň</name>
  <description>Střed města Plzeň </description>
  <Point>
    <coordinates>13.3775,49.7475,0</coordinates>
  </Point>
</Placemark>
</kml>
```

2.3.2.3 Geography Markup Language - GML

GML je formát na bázi jazyka XML, jež je od počátku vyvíjen konsorciem OGC. GML je učené pro prostorová data, je spojené především s ostatními

standardsy OGC jako je např. Web Feature Service. GML je taktéž používaný formát v datech podléhajících směrnici INSPIRE. GML je formát, v němž jsou dle směrnice poskytovány data příslušných orgánů státní správy (např. katastru nemovitosti). Vhodný je především pro komplexní data, jež se týkají směrnice INSPIRE. Pro jednoduchá data není vhodný, protože jeho využitelnost v GIS aplikacích není vždy zcela bez problémů.

Ukázka GML:

```
<gml:Point gml:id="p21"
srsName="http://www.opengis.net/def/crs/EPSG/0/4326">
  <gml:pos srsDimension="2">13.3775,49.7475,0</gml:pos>
</gml:Point>
```

2.3.2.4 ShapeFile SHP

Formát SHP je binárním formátem na základě otevřeného standardu, jež je vyvinut firmou ESRI. Ačkoliv ESRI není nezávislá organizace, je formát SHP široce rozšířen a v GIS komunitě všeobecně používaným. Jeho použití ve vhodné především tam, kde se očekává využití dat v GIS aplikacích a jejich další analýze.

2.3.2.5 GeoJSON

GeoJSON je formát postavený na jazyce JSON a je standardizován OGC. Jeho využití je především ve webových aplikacích.

```
{ "type": "FeatureCollection",
  "features": [
    { "type": "Feature",
      "geometry": {"type": "Point", "coordinates":
        [13.3775,49.7475,0]},
      "properties": {"nazev": "Plzen"}
    }
  ]
}
```

2.3.3 Rastrová data

Rastrová data jsou proti vektorovým poměrně jednoduchá. K dispozici je řada formátů popsaných pomocí specifikací OGC. V GIS je rozšířen zejména GeoTiff, případně JPG. Klíčové rozšíření oproti standardním rastrovým formátům je doplněné tzv. georeferenci. Například v případě JPG je to tzv. WorldFile a PRJ soubor vyjadřující kartografické zobrazení a afinní transformaci rastru do geoprostoru.

2.3.4 Otevřené datové služby

V případě, že velikost dat neumožňuje nabízet data ke stažení po souborech, je vhodné poskytnout data pomocí webových služeb. Nejdůležitější služby s otevřenou specifikací v oblasti prostorový data jsou:

- WMS - Web Map Service - služba určená pro poskytování rastrů určených ke zobrazení (např. RGB).
- WFS - Web Feature Service - služba určená pro poskytování vektorových dat
- WCS - Web Coverage Service - služba vhodná k poskytování obecných prostorových dat většinou rastrového charakteru (např. data o geofyzikálních veličinách).

Z hlediska poskytovatele i uživatele dat je vhodné použít řešení, které je nejjednodušší možné pro daný účel. Např. použití komplexního GML je žádoucí pouze tehdy, pokud skutečně není možné data dostatečně dobře uchopit jiným způsobem. Obdobně RDF je vhodné použít pouze pokud se očekává využití dat např. pro jazyk SPARQL. Na základě zkušeností z prostředí GIS také platí, že ačkoliv formát SHP není standardizován nezávislou organizací, je jeho použití vhodné a velice snadno využitelné v prostředí GIS.

V případě potřeby poskytování dat prostřednictvím služeb (velkoobjemová data, často aktualizována data) se doporučuje využití standardů OGC (zejména WMS, WFS a WCS). Důvodem je jejich návaznost na direktivu INSPIRE. Problematikou služeb a formátů se zabývá i národní GeoInfoStrategie, zejména v rámci v projektu TB9500MV002 Vytvoření národních technických specifikací služeb nad prostorovými daty a metadaty. Tato problematika je

dále náplní následujících opatření GeoInfoStrategie – O15 Zpracování analýzy prostorových informací a principů jejich správy v informačních systémech veřejné správy, O16 Připravit a publikovat klíčové sady prostorových dat jako otevřená data a O76 Implementace směrnice INSPIRE.

2.4 Publikování datových modelů

Datovým modelem se rozumí popis struktury dat a vzájemné vztahy jednotlivých datových prvků. V současnosti lze na datové modelování nahlížet v několika úrovních:

- Konceptuální datový model - nejvyšší míra abstrakce, model obsahuje informace, které je nutné uchovávat. Konceptuální model charakterizuje, CO uchovávat.
- Technologický (logický) datový model - v technickém modelu je zohledňováno technologické řešení problému. Model definuje, JAK budou uchovávaná data strukturována (zda bude použit relační model, objektový, síťový apod.)
- Implementační model - Model na nejnižší úrovni kde je řešena vlastní implementace v konkrétním systému (např SQL).

Pro prostorová data a v souvislosti s otevřenými daty je stěžejní Technologický (datový model), kterému se věnujeme v této části. Datové modely lze chápat jako samostatný datový zdroj, nebo jako doplňkovou informaci k existujícím datům. V následujících částech uvedeme tři stupně komplexnosti datového modelu a jeho publikování.

2.4.1 Jednoduchý datový model

Pro jednoduchá data, kde je zřejmý jejich význam postačují jednoduché způsoby popisu modelu. Důležité je vysvětlit význam veličin a atributů ve vlastních datech. Ukázkou může být datová sada dopravních nehod ve Velké Británii poskytovaná na adrese <https://data.gov.uk/dataset/road-accidents-safety-data/>.

Zmíněná datová sada obsahuje tzv “Lookup tables for variables”, kde jsou ve formátu XLS k dispozici jednotlivé číselníky pro dílčí proměnné.

2.4.2 Středně složitý datový model

Data mají složitější datovou strukturu než jednoduchá tabulka. Jedná se např. o relační nebo síťový datový model, avšak jeho komplexnost je na nižší úrovni, která umožňuje jeho znázornění pomocí schématu v pro člověka srozumitelné podobě.

Příkladem může být datová sada Open Transport Map. Na stránkách <http://opentransportmap.info/> je k dispozici vizualizace datového modelu a příslušné číselníky a definice atributů.

2.4.3 Komplexní datový model

Pokud je datový model komplexní a je předpokládáno jeho softwarové zpracování (např. automatické vytvoření databázového schématu, tvorba RDF apod.) je nutné model publikovat dle relevantních standardů ve strojově čitelné podobě. Z hlediska formátů a technologií lze datový model poskytnout např. prostřednictvím následujících formátů:

- Resource Description Framework (RDF)
- Feature catalogue (ISO 19110) XML
- Unified Modeling Language (UML)

Datový model se stává důležitou součástí metadat. Způsob jeho definice a jeho implementace v metadatovém profile je podrobněji rozebrána v kapitole Aplikační schéma / struktura dat.

2.5 Identifikátory

Veřejně dostupný systém jednotných časově stálých URI tvoří základ pro publikování otevřených propojených dat na webu.

Tato metodika doporučuje veřejné správě

- vést systém jednotných identifikátorů u všech organizací nebo jejich územně oddělených složek
- stanovit systém URI mapující stávající základní registry

-
- vybudovat systém webových služeb zajišťující přístup k těmto registrům ve formě LOD.

Požadované vlastnosti:

- Dereferencovatelné URI (XML-RDF/HTML/JSON-LD).
- V RDF budou definovány možné vazby na další identifikátory (např. vazba mezi budovou a parcelou, budovou a adresou apod.).
- Služby by měly být v gesci správců základních registrů.
- Služby by měly být publikovány pod doménou data.gov.cz¹⁷.

Vytvořením těchto komponent bude stanoven referenční rámec pro vytváření dalších otevřených dat v rámci ČR.

Tvar URI a struktura RDF dokumentů by měly být předmětem širší diskuse, proto následující doporučení je nutno brát jako předběžný návrh.

2.5.1 Registr osob (ROS)

Seznam právnických osob. IČ je platné pro většinu, ale některé organizace IČ nemají. Unikátní identifikace je zaručena v tomto případě např. Číslem datové schránky. Do budoucna se doporučuje zavést IČ pro všechny organizace.

Dočasné řešení: V případě, že neexistuje IČ, je možno použít k identifikaci číslo datové schránky.

Tvar URI (návrhy)

`http://data.gov.cz/reg/ROS/[IC/<IČ-organizace>|DS/<číslo-dat-schránky>]`

Nebo například

`http://data.gov.cz/reg/ROS/[<IČ-organizace>|DS-<číslo-dat-schránky>]`

Příklad (IČO): `http://data.gov.cz/reg/ROS/00007064`

¹⁷Pro ilustraci komplikovanosti problematiky definování URI autoři připojují komentář Doc. Mgr. Martina Nečaského, Ph.D., který publikoval ve veřejné pracovní verzi tohoto dokumentu: „Po poměrně intenzivní diskuzi s linked open data a sémantickými kolegy jsme došli k závěru, že ta URI by neměla být centralizována na doméně gov.cz, ale na doméně příslušného poskytovatele. Tj. v případě RUIAN na doméně cuzk.cz nebo ruian.cz. Už jen proto, že URI musí být dereferencovatelná a poskytovatel k nim musí poskytnout relevantní údaje. Jak je má při dereferenci poskytnout, když nemá doménu pod svojí správou.“

Příklad (Datová schránka): <http://data.gov.cz/reg/ROS/DS-123456>

2.5.2 Registr obyvatel (ROB)

Jedná se o důvěrná data, ale mohla by existovat služba založená na podobných principech jako organizace (Neveřejná).

2.5.3 Registr územní identifikace, adres a nemovitostí (RÚIAN)

Tento registr je nejvýznamnější pro oblast geoinformatiky.

Sledované entity:

- Stát
- Region soudržnosti
- Kraj (VÚSC)
- Okres
- ORP
- POU
- Obec, vojenský újezd
- Správní obvod Prahy
- Městská část/obvod
- Část obce
- Volební okrsek
- Ulice
- Stavební objekt
- Adresní místo
- Katastrální území
- Parcela
- ZSJ

Doporučení:

- V URI použít zkratky datových témat INSPIRE, pokud jsou definovány.
- Použít v maximální míře stávající číselníky a kódy používané ČÚZK.

Příklad - adresní místo: <http://data.gov.cz/reg/RUAIN/AD/123456>

2.5.4 Registr práv a povinností (RPP)

Registr práv a povinností slouží jako zdroj údajů pro informační systémy základních registrů při řízení přístupu uživatelů k údajům v jednotlivých registrech a agendových informačních systémech. Obsahem jsou

- referenční údaje o působnosti orgánů veřejné moci, a to o agendách, o orgánech veřejné moci, které je vykonávají, o informačních systémech, které pro výkon agend používají a o rozsahu oprávnění přístupu k referenčním údajům a návaznost na eSbírku,
- referenční údaje o právech a povinnostech osob, a to o údajích o rozhodnutích, na jejichž základě dochází ke změně referenčních údajů v základních registrech, o údajích o dalších právech a povinnostech osob, pokud tak stanoví jiný právní předpis.

Tento registr by měl být zdrojem informací o vazbách mezi pořizovanými daty a odpovídající agendou ve veřejné správě.

Příklad: <http://data.gov.cz/reg/RPP/123456>

Poznámka: V současné době jsou data RPP přístupná na adrese: <https://rpp-ais.egon.gov.cz/gen/agendy-detail>, tedy pro konkrétní agendu například <https://rpp-ais.egon.gov.cz/gen/agendy-detail/Agenda123.html>. Takto je navrhováno použití v kapitole o metadatech.

2.6 Vazby na externí datové zdroje

Nejvyšší úroveň Linked Open Data je spojena s vazbami na externí data a slovníky. Tyto vazby plní tzv. sociální funkci LOD, protože napomáhají k vytváření datových struktur napříč různými organizacemi, státy a kulturami. V případě prostorových dat můžeme typy vazeb rozdělit do tří skupin:

1. Identické a podobnostní vazby
2. Topologické vazby
3. Hierarchické a mereologické vazby

2.6.1 Identické a podobnostní vazby

Tato kapitola vychází především ze standardů OWL (W3C 2012, Dean et al. (2004)), SKOS (Miles a Bechhofer 2009) a dalších. Následující přehled ukazuje, jak jsou definované jednotlivé vlastnosti propojující identické (shodné, ekvivalentní) a podobné objekty. Seznam vlastností je abecední, přičemž řazení i popis relace využívá i standardní prefixy (například **skos:** nebo **owl:**).

OWL disponuje i opačnou vlastností **owl:differentFrom**, která by se v případě prostorových hodila pro odlišení různých prvků se stejným zeměpisným jménem.

2.6.1.1 owl:sameAs

Vlastnost **owl:sameAs** je speciálním případem relace **rdf:Property**. Podle Dean et al. (2004) vlastnost propojuje dva individuály (tedy nikoli třídy¹⁸), proto se často využívá k mapování mezi ontologiemi. Referenční příručka (Dean et al. 2004) tvrdí, že vlastnost **owl:sameAs** poukazuje na to, „že dva identifikátory (URI) odkazují na stejný objekt, tyto individuály mají stejnou identitu“. Na tomto místě je potřeba poznamenat, že v originální definici je poslední slovo předchozího citátu, tedy „identita“ psáno v uvozovkách, což zcela jistě znamená, že se jedná o poměrně vágní pojem, především ve světě komunikace a přenosu informací.

2.6.1.2 rdfs:seeAlso

V některých řídkých případech (například GeoNames.org) se pro odkazy na příbuzné (tedy i ekvivalentní a podobné) prvky používá i tato velice obecná vlastnost standardizovaná v rámci RDFS (Brickley a Guha 2014). Vlastnost **rdfs:seeAlso** má podle standardu Brickley a Guha (2014) „poukazovat na zdroj, který může poskytovat doplňkové informace o subjektu“. Při výzkumu identických a podobných objektů je nutné v případě použití vlastnosti **rdfs:seeAlso** přesvědčit se o charakteru datového zdroje a způsobu použití výše jmenované vlastnosti.

¹⁸Výjimkou je dialekt OWL Full, kde se třídami zachází stejně jako s instancemi, a vlastnost **owl:sameAs** může tudíž definovat ekvivalenci tříd. Pro propojení dvou stejných tříd se obvykle používá vlastnost **owl:equivalentClass**.

2.6.1.3 skos:closeMatch

Miles a Bechhofer (2009) uvádí, že vlastnost **skos:closeMatch** „je používána k propojení dvou konceptů, které jsou dostatečně podobné, takže mohou být zaměnitelně použity v aplikacích pro vyhledávání informací“. Vlastnost **skos:closeMatch** není definována jako tranzitivní (především z důvodu toho, že „dostatečná podobnost“ je velice vágní a neurčité kritérium pro propojování konceptů napříč znalostními bázemi).

2.6.1.4 skos:exactMatch

Podle Referenční příručky standardu SKOS (Miles a Bechhofer 2009) vlastnost **skos:exactMatch** je použita „k propojení dvou konceptů vykazujících vysoký stupeň důvěry, že tyto koncepty mohou být využívány nezaměnitelně ve velkém množství aplikací pro vyhledávání informací“. Vlastnost **skos:exactMatch** je tranzitivní a je pod-vlastností **skos:closeMatch**.

Standard SKOS (Miles a Bechhofer 2009) ještě uvádí další vlastnosti, které slouží ke vzájemnému mapování zdrojů. Konkrétně se jedná o relace **skos:broadMatch**, **skos:narrowMatch** a **skos:relatedMatch**. V tomto dokumentu však nejsou více rozváděny, protože jejich využití je sporé a nevyskytují se ani v případových studiích.

Kromě čtyř výše uvedených standardizovaných vlastností pro zápis vazeb mezi identickými nebo podobnými prvky používají různé znalostní báze a další podobné nástroje využívající princip Linked Data je další vazby, které můžeme označit jako produktově specifické. Mezi nejdůležitější patří Wikidata Identifiers. V této sekci se vyskytují položky, které reprezentují odkazy na ekvivalentní nebo vysoce podobné objekty v jiných datových sadách. Problémem je, že každý externí zdroj má definovanou specifickou vlastnost. Například vlastnost P15663 představuje identifikátor v GeoNames.org.

Z předchozích řádků jsou evidentní následující závěry a doporučení:

1. V oblasti Linked Open Data neexistuje jedna dominantní vlastnost propojující stejné a podobné prvky.

-
2. Výše uvedené definice vlastností obsahují řadu vágních pojmů, které znesnadňují rozhodování, kdy jsou objekty skutečně identické nebo podobné.
 3. Při volbě vhodné identické a podobnostní vlastnosti by měly být upřednostňovány existující standardy (SKOS, OWL, RDFS).
 4. Při rozhodování, kterou z výše uvedených standardizovaných vlastností zvolit, by se uživatel měl orientovat podle „síly vyjádření ekvivalence a podobnosti“. Jako nejobecnější vlastnost byla shledána `rdfs:seeAlso` (kde jsou podmínky ekvivalence a podobnosti vyjádřeny nejvíce obecně a vágně; tato vlastnost se často používá k propojení dat na nestrukturované prvky týkající se konkrétního objektu, například na fotografie, dokumenty apod.). Naopak `owl:sameAs` je nejstriktněji definovaná.
 5. Vlastnost `owl:sameAs` by měla propojovat pouze třídy nebo instance, zatímco `rdfs:seeAlso` je vlastnost anotační a vlastnosti `skos:exactMatch`, případně `skos:closeMatch` mohou být také využity jako anotační, což zvyšuje jejich využitelnost.

Příklady využívání identických a podobnostních vlastností jsou uvedeny v modelové aplikaci Smart Points of Interest.

2.6.2 Topologické vazby

Topologické vazby (vztahy, pravidla) definují propojení mezi jednotlivými objekty na základě geometrických principů. Proto je tento typ vazeb specifický pro prostorová data, která často vyjadřují polohu nebo vymezení prvků nikoli pomocí nepřímého geokódování.

Topologické vazby jsou důležité i z pohledu dotazovacích jazyků. Kdy jazyky jako GeoSPARQL dokáží definovat vazby na bázi topologických vztahů (například „najdi objekty ležící ve městě“ apod.).

Vlastní topologická pravidla jsou definována a popsána v mnoha publikacích, například Longley et al. (2001). Tato metodika se zabývá přepisem těchto pravidel do RDF trojic pomocí standardizovaných slovníků. Topologické vazby jsou formalizovány pomocí několika modelů jako například Region Connection Calculus (RCC), Simple Features Topological Relations nebo Egenhofer nine-way intersection model.

Open Geospatial Consortium nadefinovalo přepis slovníku RCC8 (8 topolo-

gických vlastností zaměřené především na vztahy mezi polygony) do podoby sémantického slovníku, kde jednotlivé vlastnosti jsou identifikovány pomocí URI. Následující tabulka (Tabulka 2) obsahuje kromě URI vlastností (jmenný prostor s prefixem geo:¹⁹ a označení vlastnosti ve slovníku), anglický a česká název vlastnosti.

Tabulka 2: Slovník RCC8.

Vlastnost (česky)	Vlastnost (anglicky)	OGC URI
Identické objekty	equals	geo:rcc8-ec
Objekty, které nejsou spojeny, ani se nedotýkají	disconnected	geo:rcc8-dc
Dotýkající se objekty	externally connected	geo:rcc8-ec
Částečně se překrývající objekty	partially overlapping	geo:rcc8-po
Jeden objekt je podmnožinou druhého	non-tangential proper part	geo:rcc8-ntpp
Inverzní vlastnost k předchozí	non-tangential proper part inverse	geo:rcc8-ntppi
Jeden objekt je podmnožinou druhého a zároveň se oba objekty dotýkají	tangential proper part	geo:rcc8-tp
Inverzní vlastnost k předchozí	tangential proper part inverse	geo:rcc8-tpi

Standard GeoSPARQL (Perry a Herring 2012) obsahuje v kapitole 7 transkripce obou dalších výše jmenovaných slovníků do podoby Linked Data slovníků. Dále v tomtéž dokumentu jsou v příloze B2 příklady dotazů s topologickými vazbami.

Ukázky používání topologických vazeb jsou k dispozici například v modelové aplikaci Smart Points of Interest, kde je použitý slovník Simple Features Relation Family.

¹⁹<http://www.opengis.net/ont/OGC-GeoSPARQL/1.0/>

Pro používání topologických vazeb platí podobná pravidla jako v předchozích i následujících případech typů vlastností:

1. Uživatel by si měl vybrat model, který chce (například na základě předchozí zkušenosti nebo implementace v software) používat.
2. V první řadě by měly být objekty propojené pomocí vazeb standardizovaných v modelech.
3. Pokud se v modelu (resp. jeho formalizaci do podoby Linked Data slovníku) daná vlastnost přímo nevyskytuje, je žádoucí ji do takové vazby transformovat (například vlastnost “leží v sousedství” může být nahrazena vlastností “dotýká se”).
4. Nepodaří-li se předchozí krok, pak je vhodné poohlédnout se po jiném řešení, které už podobnou vlast použilo a tedy i definovalo. Příkladem takového řešení mohou být aktivity Center of Excellence for Geospatial Information Science (CEGIS), především Ontology for The National Map²⁰.
5. Další možností je nahrazení topologické vazby standardizovanou hierarchickou nebo mereologickou vlastností (viz níže).
6. Poslední, nejméně vhodnou, možností, protože tato varianta snižuje interoperabilitu, je definování si vlastní topologické vazby.

2.6.3 Hierarchické a mereologické vazby

Tyto vazby nejsou primárně určeny k propojení na jiná externí data, ale především ke klasifikaci a vnitřnímu rozdělení prvků v datové sadě. Vzhledem k principu LOD je však zřejmé, že objektem (cílem) relace nemusí být nutně prvek ve stejné datové sadě, ale může se jednat o libovolnou identitu používající URI jako identifikátor.

Nejprve je vhodné uvést rozdíl mezi hierarchickou a mereologickou vlastností.

- Hierarchická vlastnost je založena na specifičnosti a obecnosti. Příkladem může být vazba mezi objekty “poušť” a “písečná poušť”, kdy “písečná poušť” je specifickým případem entity “poušť”, která je naopak obecným případem prvku “písečná poušť”. Ontologie (jako jeden z

²⁰<http://cegis.usgs.gov/ontology.html>

případů LOD), které s hierarchií často pracují, používají označení nadřazená a podřazená třída (superclass a subclass). Existuje také možnost, kdy se nejnižší stupeň hierarchie dále nečlení. Pak je v rámci ontologií popisován jako instance (individuál).

- Mereologické vlastnosti se používají k propojení částí a celku. Aplikujeme-li mereologickou vlastnost na příklad pouští, pak je možné říct, že prvky “zrnko písku” nebo “oáza” jsou částmi (nikoli specifickými případy) objektu “poušť”. Na rozdíl od hierarchických vlastností vlastnosti část-celek nejsou standardizovány (nebo se konkrétní relace uvedené ve standardech příliš nepoužívají v praxi). Proto se následující odstavce (s výjimkou závěrečných doporučení) budou týkat především hierarchických vlastností.

Jak již bylo uvedeno výše hierarchické vlastnosti jsou definovány zpravidla v ontologických strukturách. Druhou možností jejich využití jsou tezaury. Pro jejich tvorbu je určený standard SKOS, který pracuje se dvěma základními relacemi, které popisují strukturu tezauru:

1. Vazba na obecnější pojmy – **skos:broader** (existuje i tranzitivní varianta této vazby **skos:broaderTransitive**).
2. Vazba na specifitější termíny – **skos:narrower** (případně **skos:narrowerTransitive**).

Standard OWL pro realizaci hierarchických struktur dvě základní objektové vlastnosti:

1. **rdfs:subClassOf** pro definování podtříd.
2. **owl:NamedIndividual** ve spojení s **rdf:type** (odkaz na třídu nadřazenou dané instanci) pro relaci mezi individuálem a třídou.

Doporučení:

- Používají-li se v datech hierarchické i mereologické vazby, není možné je kombinovat a popisovat pomocí stejné vlastnosti. Na tomto místě je zapotřebí uvést, že v tomto bodě se teorie často liší od praxe. V mnoha tezaurech jsou vlastnosti **skos:narrower** a **skos:broader** používány pro definování obou typů vazeb.
- Podobně jako v předchozích případech je žádoucí držet se standardů, které jsou v rámci hierarchických vazeb definovány v rámci jazyků

OWL a SKOS.

- Podobně jako v případě ekvivalentních a podobnostních vazeb i hierarchické relace mají odlišný význam. Zatímco v případě standardu OWL se jedná o striktní vymezení instancí podtříd a nadřazených tříd, u standardu SKOS (přístup vhodný pro tezaury) jsou vazby **skos:narrower** a **skos:broader** mnohem vágnější.
- U mereologických vazeb je na prvním místě především jejich implicitní sémantika. Proto se doporučuje používat pro pojmenovávání takových vlastností pojmy jako “is Part Of”, “is Component Of” apod.

Kapitola 3

Modelové aplikace

3.1 Metadata

3.1.1 MICKA

Micka (Obrázek 1) je systém pro správu metadat, který podporuje CSW 2.0.2 - ISO AP 1.0, GeoDCAT, OpenSearch, a další standardy.

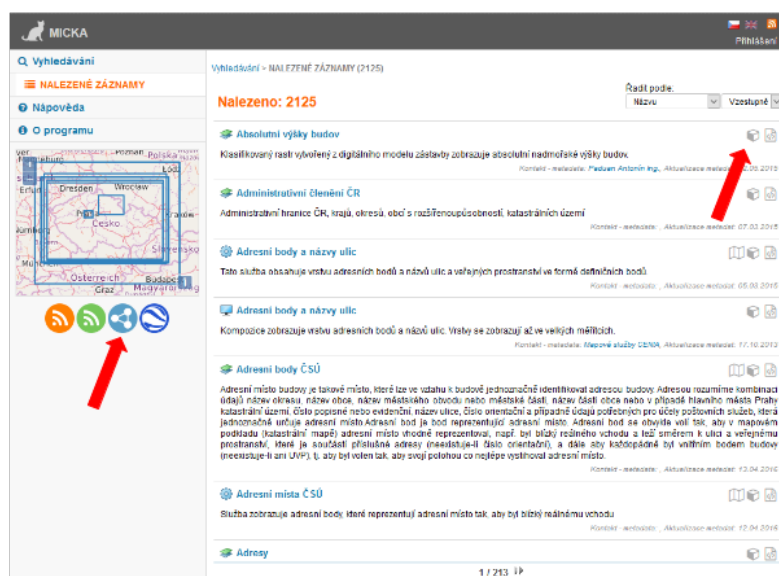
Výstup GeoDCAT je jednak umožněn pro každý metadatový záznam, jednak jako feed všech nalezených záznamů vybraných podle určitých kritérií.

Funkčnost je možno vyzkoušet na adrese: <http://geoportal.gov.cz/php/micka5>.

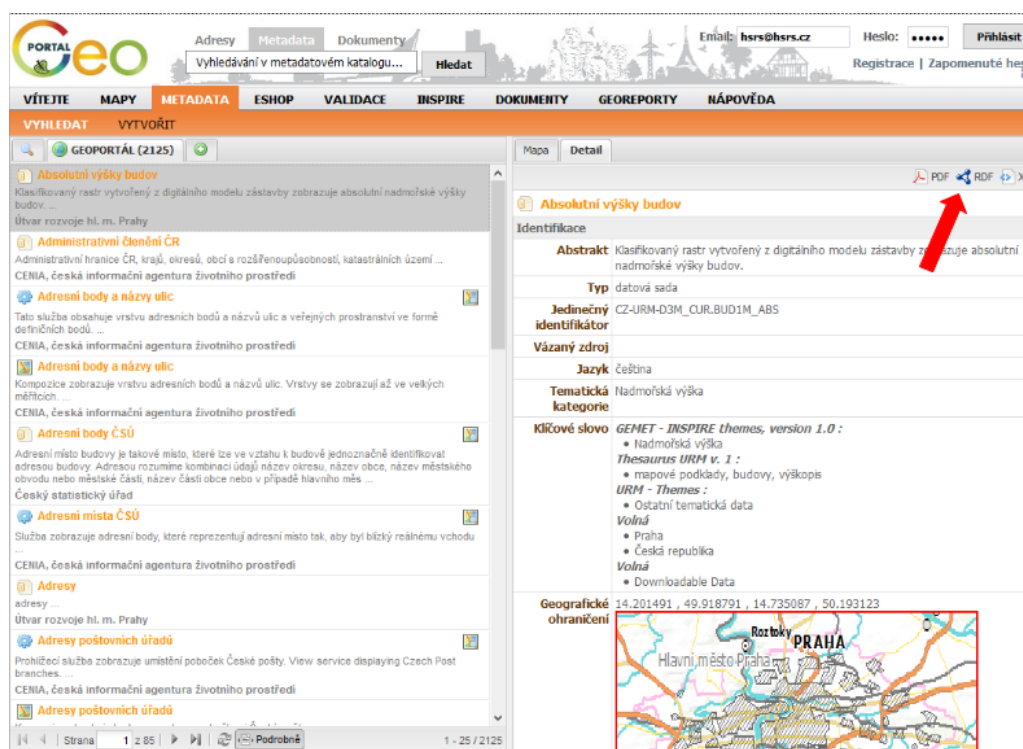
3.1.2 Katalogový klient na Národním geoportálu INSPIRE

Katalogový klient zobrazuje metadata uložená v národním katalogu INSPIRE (Obrázek 2), ale umožňuje připojit další katalogy pracující s CSW 2.0.2. Součástí nabízených funkcí je export jednotlivých záznamů v GeoDCAT.

Funkčnost je možno vyzkoušet na adrese: <https://geoportal.gov.cz/web/guest/catalogue-client>.



Obrázek 1: Uživatelské rozhraní katalogu MICKA. Červené šipky označují možnost exportu GeoDCAT. Vlevo - feed, vpravo jednotlivý záznam.



Obrázek 2: Uživatelské rozhraní katalogového klienta. Červená šipka označuje možnost exportu jednoho metadatového záznamu do GeoDCAT.

3.1.3 On-line převodník

Převodník převádí jednotlivé záznamy ISO 19139 do GeoDCAT. Předpokládá se v budoucnu jeho integrace do národního geoportálu INSPIRE. Zadejte buď URL metadatového záznamu nebo nahrajte soubor z lokálního počítače (Obrázek 3).

Je k dispozici zatím na adrese: <https://geoportal.gov.cz/php/catalogue/convert/>.

Převod metadat z INSPIRE na GeoDCAT

The screenshot shows a web form titled "Převod metadat z INSPIRE na GeoDCAT". It contains three input fields: "URL" with a text box, "Soubor" with a "Procházet..." button and the text "Soubor nevybrán.", and "Výstup" with a dropdown menu set to "DCAT". At the bottom are two buttons: "Převést" and "Vyčistit".

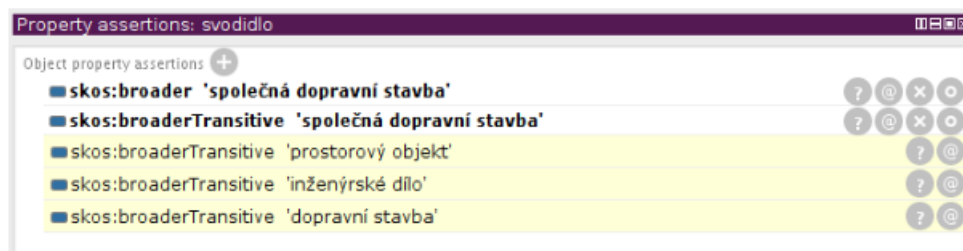
Obrázek 3: Zkušební rozhraní převodníku.

3.2 Ontologie VF XML

Ontologie VF XML je součástí projektu vytvoření výměnného formátu digitální technické mapy. Jejím primární účelem je evidence, klasifikace a popis objektů zobrazovaných v DTM a případně v dalších datových sadách. Tato modelová aplikace tedy představuje návrh slovníku, který by měl sloužit pro sémantickou podporu otevřených prostorových dat a tak snížit (nebo zcela potlačit) neurčitosti a vágnosti vznikající při interpretaci a dalším používání dat.

Při tvorbě ontologie bylo využíváno přístupu pro tvorbu tezaurů, proto ontologie využívá základní hierarchické vlastnosti formátu SKOS (Simple Knowledge Organization System), které definují vazby mezi jednotlivými typy objektů (Obrázek 4). Konkrétně se jedná o relace

- `skos:broader` - odkaz na typ objektu se širším významem (například typ objektu `stavba` je pomocí této vlastnosti spojený s typem objektu `prostorový objekt`).
- `skos:broaderTransitive` - tranzitivní varianta předchozí vlastnosti, která dokáže propojit i typy objektů v různých úrovních (například jakýkoli typ objektu s kořenovým prvkem `prostorový objekt`).
- `skos:narrower` - odkaz na typ objektu s užším významem (například typ objektu `stavba` je pomocí této vlastnosti spojený s typem objektu `pozemní stavba`).
- `skos:narrowerTransitive` - tranzitivní varianta předchozí vlastnosti.

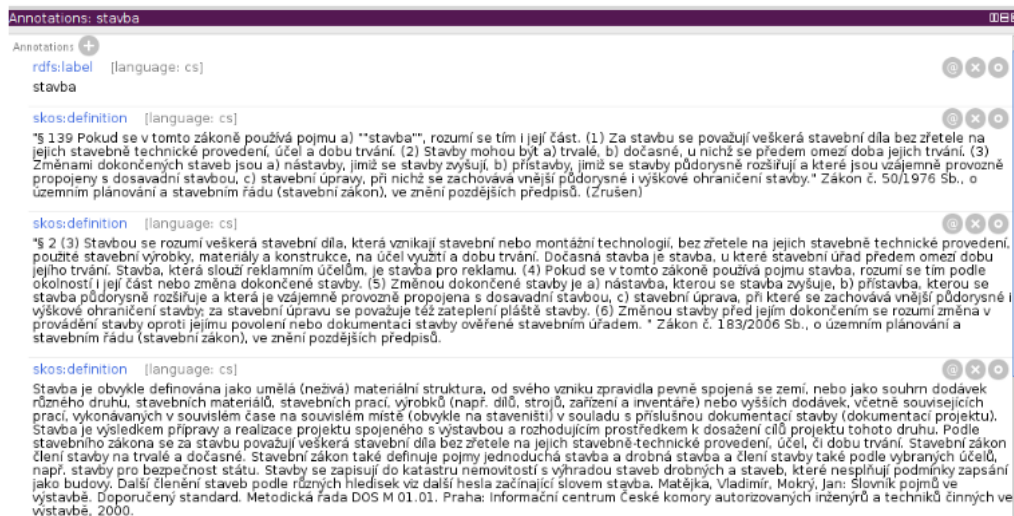


Obrázek 4: Ukázka objektových relací Ontologie VF XML.

Kromě těchto objektových vlastností jsou ze SKOS použity také anotační vlastnosti `skos:definition` a `skos:note` (Obrázek 5). Tyto vlastnosti detailněji popisují typy objektů, včetně odkazů na příslušné normy a standardy.

Ontologie také obsahuje vlastnosti umožňující zápis typu geometrie a atributů. Tímto způsobem bude možné definovat konkrétní prvky výměnného formátu. Například `hranice budovy` jako prvek výměnného formátu (tedy element prostorových dat) sice nebude popsána ontologií, protože nejde o samostatný objekt, avšak vazba mezi výměnným formátem a ontologií bude realizována odkazem na kombinace typu geometrie (v tomto případě `linie`) a typu objektu (`budova`).

Rámcový ontologický popis objektů digitální technické mapy vychází z Klasifikace stavebních děl CZ-CC (dále jen „Klasifikace CZ-CC“) vypracované Českým statistickým úřadem na základě mezinárodního standardu Classification of Types of Constructions – CC, vydaného Eurostatem v říjnu 1997. Za základ při vypracování klasifikace CC byl vzat oddíl 52 mezinárodní klasifikace produkce - CPC (Central Product Classification), kterou publikovala



Obrázek 5: Ukázka anotačních vlastností Ontologie VF XML.

Statistická komise OSN v roce 1991. Ve srovnání se strukturou CPC je v klasifikaci CC sekce „Budovy“ rozdělena podrobněji, aby dostatečně zahrнула potřebný počet klasifikačních úrovní (stupňů).

Klasifikace je sestavena tak, aby mohla být využita například i pro statistiku stavebních činností, sčítání domů a bytů, statistiku cen stavebních prací nebo pro národní účetnictví. Klasifikace je používána pro definici staveb, která je potřebná pro zabezpečení informací o specifických proměnných krátkodobých ukazatelích (např. stavební povolení, kolaudace, uvedení do provozu). Lze ji rovněž využívat v průběhu celé životnosti stavby pro evidenci změn užívání, pro obchodní jednání, opravy, rekonstrukci a modernizaci. Účelné může být rovněž její využití jako klasifikačního standardu pro vyhlašování tendrů nebo uzavírání kontraktů v rámci zadávání veřejných zakázek, využita může být pro potřeby daňové a rozpočtové. Uvažuje-li se v budoucnu při údržbě a vedení DTM s využitím geodetické části dokumentace skutečného provedení staveb dílčích investičních celků, představuje tento přístup jednotné systémové řešení pro komplex činností od podkladů pro projekt, po celou dobu jeho životnosti, až po proces zániku daného objektu.

Prostorové objekty jako nejobecnější objekty jsou děleny na stavby, inženýrská díla, přírodní a polopřírodní objekty, abstraktní objekty a byla doplněna

kategorie odvozený objekt. Stavby se klasifikují podle technického řešení stavby (projektu), které vyplývá ze zvláštního užívání stavby (např. budovy pro obchod, komunikace, díla vodní, vedení dálková trubní apod.); budovy jsou klasifikovány podle jejich hlavního užívání (bytové, nebytové), inženýrská díla podle projektů, které přímo určují účel a užití stavebního díla.

První úroveň ontologie (kořenový prvek) představuje univerzální **prostorový objekt**, který jako atributy obsahuje identifikátor, geometrii (typu Geometry v INSPIRE, který bude zahrnovat různé typy geometrií) a případně metadata (včetně evidence změn). Jedná se o v reálném světě existující jednotlivý objekt nebo fenomén, v modelovém světě pak o synonymum pro typ entity či geoprvek. Druhá až šestá úroveň znázorňuje hierarchické uspořádání kategorií základních objektových typů ontologie, včetně pojmenování základního objektového typu ontologie na nejnižší vedené úrovni daného objektu.

Rámcový ontologický popis je navržen pro možnost univerzálního rozšíření na jakékoli úrovni jeho uspořádání.

Na závěr je nutné poznamenat, že tvorba Ontologie VF XML není ukončena a očekávají se výrazné změny a doplnění chybějících informací. Pro základní orientaci v ontologii je možné použít vizualizaci ontologie¹ (Obrázek 6) a jejich jednotlivých prvků ve formě webových stránek. Vlastní ontologie je pak zapsána pomocí formátu OWL 2 (Web Ontology Language) s využitím všech názvných standardů (například RDFS) a serializace RDF/XML.

3.3 Smart Points of Interest

V rámci projektu SDI4apps vzniká ve spolupráci Katedry geomatiky na Západočeské univerzitě, Českého centra pro vědu a společnost a Baltic Open Solution Centre (Lotyšsko) datová sada SPOI (Smart Points of Interest). SPOI obsahuje tzv. body zájmu (points of interest), které jsou určeny především jako datový podklad pro aplikace zabývající se turistikou nebo cestovním ruchem. Informace uložené ve SPOI mohou sloužit i jako velmi účinná a efektivní propagace konkrétních památek nebo služeb.

Data nejsou uložena v tradiční relační databázi, ale jako grafová struktura – RDF (Resource Description Framework) trojice. V současné době je k dispozici

¹<http://gis.zcu.cz/projekty/Ontologie/>

chmelnice

URI (rdf:about): http://gis.zcu.cz/prostorovy_objekt#chmelnice

Popisek (rdfs:label): chmelnice

Úroveň: 4

Vazba na VF XML: chmelnice

Typy objektů se širším významem (obecnější objekty !)	Typy objektů se užším významem (specifičtější objekty !!)
Typ objektu se širším významem (skos:broader): zemědělství (3)	
Typ objektu se širším významem - tranzitivní (skos:broaderTransitive): přírodní a polopřírodní objekty (2)	
Typ objektu se širším významem - tranzitivní (skos:broaderTransitive): prostorový objekt (1)	
Typ objektu se širším významem - tranzitivní (skos:broaderTransitive): zemědělství (3)	

Zařazené atributy objektu:

datum (prostorový objekt)

provozovatel (prostorový objekt)

způsob pořízení (prostorový objekt)

id (prostorový objekt)

vlastník (prostorový objekt)

poloha (prostorový objekt)

[Zpět na seznam typů objektů](#)

Ontologie obsahuje celkem 333 objektů.

Katedra geomatiky Západočeské univerzity v Plzni, GeoReal, 2016

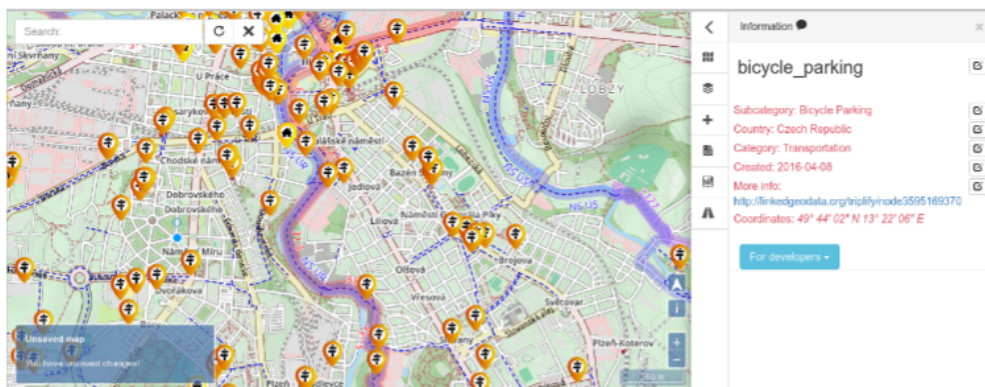
Obrázek 6: Vizualizace Ontologie VF XML.

více než 27 miliónů bodů, které může uživatel využívat při dodržení podmínek licence ODbL (Open Database License). Data je možné prohlížet v mapovém prohlížeči využívajícím knihovny HS Layers anebo stahovat pomocí SPARQL (SPARQL Protocol and RDF Query Language) endpoint ve formátech RDF, CSV, JSON a dalších.

SPOI splňuje veškeré požadavky pro tzv. pětihvězdičková Linked Open Data – otevřenost z pohledu licence, dostupnost ve strukturovaném a otevřeném formátu, jednoznačné identifikátory prvků pomocí URI (Uniform Resource Identifier) a standardizované propojení na jiná data (například na objekty v databázích Dbpedia, LinkedGeoData.org nebo GeoNames.org). Při tvorbě datové sady byla využita převážně existující data z volně dostupných zdrojů. Jedná se například o OpenStreetMap, Wikidata, Natural Earth nebo GeoNames.org. Kromě těchto globálních datových sad byly také zpracovány lokální datové sady od partnerů projektů SDI4apps (Pošumaví, Zemgale), Open Transport Net (Antverpy, Issy) a vybraná otevřená data (Praha, projekt Citadel on the Move).

Velkou výhodou řešení SPOI je kromě toho, že jde o otevřená a bezešvá data, existence jednotného datového modelu. Ten maximálně využívá existující slovníky Linked Data (například RDFS, FOAF, GeoSPARQL nebo Dublin

Core) a umožňuje propojení s dalšími objekty a slovníky. Veškeré informace (vizualizace - Obrázek 7, SPARQL endpoint, datový model, schéma datové harmonizace, aktuality...) jsou k dispozici na webové stránce sdi4apps.eu/spoi².



Obrázek 7: Mapový klient SPOI.

Ukázka RDF zápisu bodu zájmu:

```
<rdf:Description rdf:about="http://www.sdi4apps.eu/poi/#Belluno1">
<rdfs:label xml:lang="it">Sovramonte</rdfs:label>
<rdfs:comment xml:lang="it">Busa di Cavaren e Busa di Monsampiano
si trovano in comune di Sovramonte, sulle Vette Feltrine,
all'interno del Parco Nazionale delle Dolomiti Bellunesi e
all'ombra delle Grandi Vette. Entrambe sono dedicate
all'alpeggio delle vacche.</rdfs:comment>
<geos:asWKT
rdf:datatype="http://www.openlinksw.com/schemas/virttrdf#Geometry">
POINT(11.826083000000001 46.098517000000001)</geos:asWKT>
<poi:region resource="http://nuts.geovocab.org/id/ITD33"/>
<poi:class rdf:resource="http://gis.zcu.cz/SPOI/Ontology#lodging"/>
<foaf:phone>+39 331 8956681</foaf:phone>
<poi:access>"NULLA piedi: dall'abitato di Aune (a 13km da Feltre
direzione passo Croce d'Aune) s'imbocca il sentiero CAI n. 810
""Sentiero di S. Antonio"" sino alla Malga Monsampiano (3.5h),
da qui si segue il sentiero CAI 810"</poi:access>
```

²<http://sdi4apps.eu/spoi>

```

<poi:access>A piedi o in mtb: dal Passo Croce d'Aune (a 11km
nord da Feltre) tramite la mulattiera (CAI 801; 2,5h) o la strada
militare chiusa che conduce al Rif. Dal Piaz; poco prima del rifugio
seguire, sulla sinistra, il segnavia CAI n. 810</poi:access>
<poi:photo>http://www.infodolomiti.it/media/infodolomiti/Strutture/
Malghe/Malga_Cavaren_Monsampiano/Malga_Cavaren_Monsampiano_1.jpg
</poi:photo>
<rdfs:seeAlso>http://www.infodolomiti.it/dormire-e-mangiare/malghe/
busa-di-cavaren-e-busa-di-monsampiano/6972-11.html</rdfs:seeAlso>
<geos:sfWithin rdf:resource="http://dbpedia.org/resource/Italy"/>
<geos:sfWithin rdf:resource="http://www.geonames.org/3175395"/>
<dc:identifier rdf:resource="http://www.sdi4apps.eu/poi/Belluno1"/>
<dc:publisher>SPOI (http://sdi4apps.eu/spoi)</dc:publisher>
<dc:title>Sovramonte</dc:title>
<dc:rights
rdf:resource="http://opendatacommons.org/licenses/odbl/1.0/">
<dc:source rdf:resource="http://www.infodolomiti.it"/>
<dcterms:created
rdf:datatype="http://www.w3.org/2001/XMLSchema#date">
2016-08-23</dcterms:created>
</rdf:Description>

```

Z hlediska otevřených dat jsou důležité především vazby na jiné datové zdroje a slovníky. SPOI využívá následující vlastnosti k propojení na externí data:

- **foaf:homepage** - propojení na webovou stránku bodu zájmu. Vlastnost je definována ve standardu FOAF (Friend of a friend). Formát FOAF obsahuje i další vlastnosti. Ve SPOI jsou použity ještě **foaf:phone** (zápis telefonního čísla) a **foaf:mbox** (zápis emailové adresy).
- **rdfs:seeAlso** - link na webové stránky popisující POI (například články na Wikipedii, Wolfram|Alpha nebo mapy a plány ve formě rastrů).
- **skos:exactMatch** a **owl:sameAs** - propojení na identické prvky v jiných datových sadách. Z důvodu vyšší interoperability jsou zatím použity oba atributy, protože tento způsob zvyšuje průchodnost topologického grafu zobrazujícího podobné nebo ekvivalentní objekty v různých datových sadách.
- **geos:sfWithin** - topologická vazba podle standardu GeoSPARQL vycházející z přístupu Simple Features. Ve SPOI je tato vazba využita pro propojení objektu s příslušným státem. Ten je uveden pomocí URI v

databázích GeoNames.org a DBpedia.

Pro každý objekt SPOI jsou publikována stručná feature metadata vycházející ze standardu Dublin Core. Konkrétně se jedná o identifikátor, licenci, název prvku, datum zařazení do databáze a název datové sady.

Souřadnice v systému WGS 84 jsou zapsány pomocí WKT serializace (zeměpisná délka zeměpisná šířka) v GeoSPARQL standardu. Z důvodu lepší kompatibility se systémem Virtuoso, který zajišťuje databázové uložení dat, byl původní atribut `poi:hasExactGeometry` propojen na datový typ `sf:Point` nahrazený vlastností `geos:asWKT`.

3.4 Open Transport Map

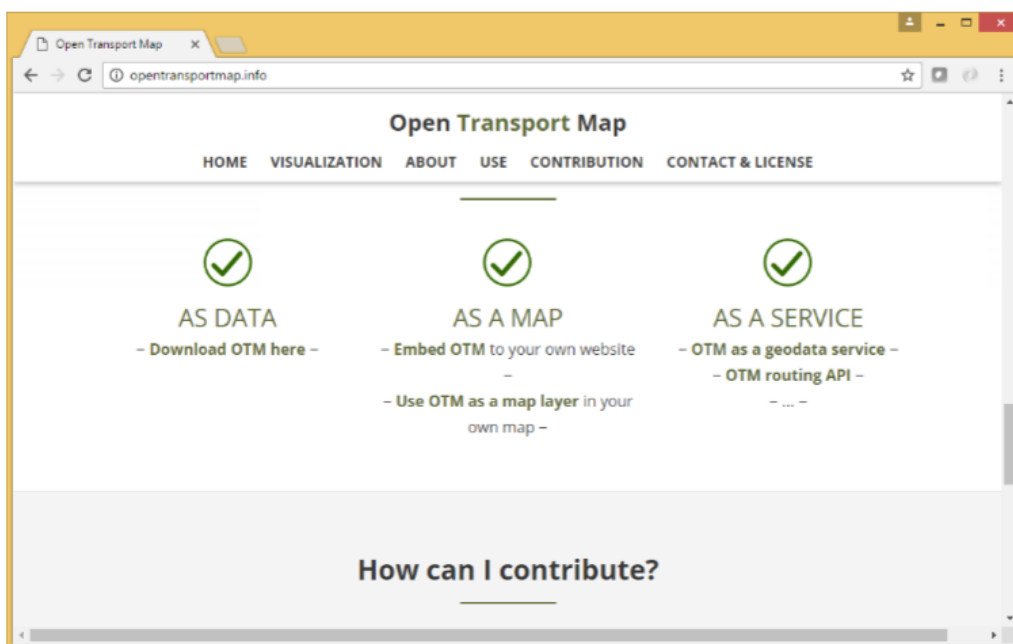
Open Transport Map (česky Otevřená dopravní mapa) je příkladem datové sady poskytující prostorová data o pozemních komunikacích pro území Evropy ve formě otevřených dat. Datová sada je k dispozici na adrese: <http://opentransportmap.info/>.

Hlavní vlastnosti této datové sady jsou:

- Datový model kompatibilní s INSPIRE.
- Datový model použitelný pro hledání nejkratší cesty (routing).
- Datový model rozšířený pro informace o intenzitě dopravy.
- Jedná se o rozsáhlou datovou sadu (cca 50 GB).
- Jedná se o liniová vektorová data.
- Datová sada je vydávána pod licencí Open Data Commons Open Database License (ODbL).

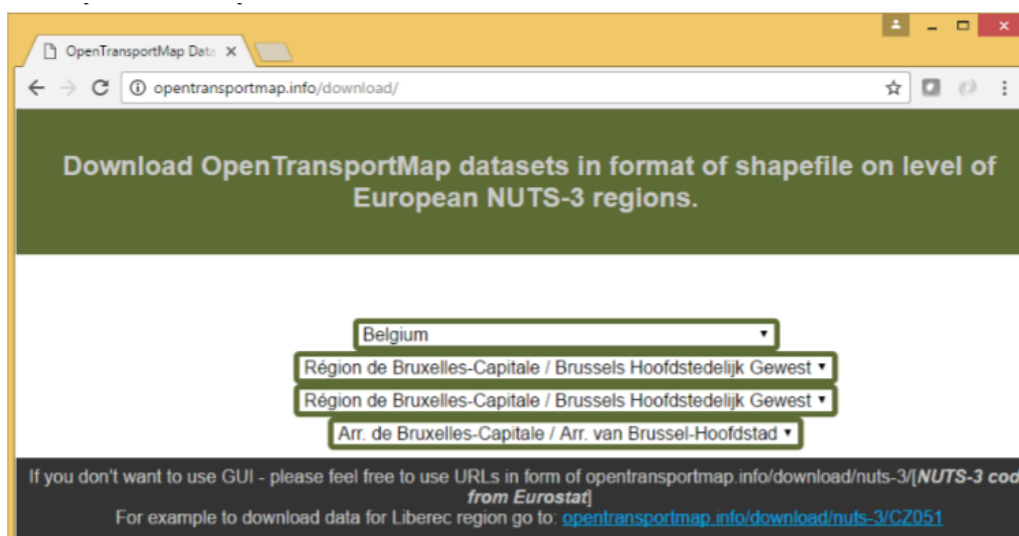
Datová sada je odvozena především z dat OpenStreetMap a byla vytvořena v rámci projektu OTN.

Z hlediska metodiky publikování otevřených dat se jedná o prostorovou datovou sadu, jež je k dispozici v několika různých podobách (Obrázek 8), které odpovídají různým úrovním pětihvězdičkového modelu. Důvodem je velikost datové sady (dopravní síť celé Evropy), která neumožňuje přímočarý způsob poskytnutí dat ve formě jednoho souboru. K dispozici jsou následující postupy:



Obrázek 8: Ukázka možnosti přístupu k datům OTM.

- As data - V tomto případě je možné stáhnout jednotlivé části dat v proprietárním (avšak v geo-komunitě velice rozšířeném) formátu SHP. Při stažení je dostupné i permanentní URL příslušné datové sady. Tuto službu lze hodnotit jako dvouhvězdičkovou (Obrázek 9), protože data jsou strojově čitelná, avšak nikoliv v neproprietárním formátu.
- As a Map - Zde jsou data dostupná v rastrové podobě pomocí formátu standardu WMS (označováno také jako prohlížečská služba). Výhodou je nabízená aplikace, která umožňuje přímé začlenění datové sady do konkrétní webové stránky. Tuto službu lze hodnotit jako jednohvězdičkovou, protože neumožňuje strojové čtení dat (jedná se vlastně o vektorová data, poskytovaná formou rastru). V případě jiného typu dat (např. ortofoto) by však tato metoda splňovala minimálně tři hvězdičky.
- As a service - datová sada ve formě WFS (stahovací služba) a ve formě Linked Open Data (RDF). Tato úroveň zatím není na webových stránkách k dispozici. Metoda by splňovala tři hvězdičky a více. Postup použitý pro tvorbu RDF je popsán níže.



Obrázek 9: OTM jako dvouhvězdičková linked open data.

3.4.1 Datový model OTM

Pro publikování dat je důležitá nejen forma jejich poskytování, ale rovněž datový model. Pro účely OTM došlo k vytvoření datového modelu (viz Obrázek 10), který je kompatibilní se směrnicí INSPIRE a používá příslušné číselníky a vztahy mezi prvky. Číselníky pro jednotlivé kategorie (viz Obrázek 11). K tomuto účelu bylo vytvořeno v rámci projektu OTN konverzní schéma, které definovalo vztahy mezi vstupním datovým modelem (většinou OpenStreetMap) a uvedeným cílovým datovým modelem.

Ukázka konverzního schématu (Tabulka 3):

«codeList»

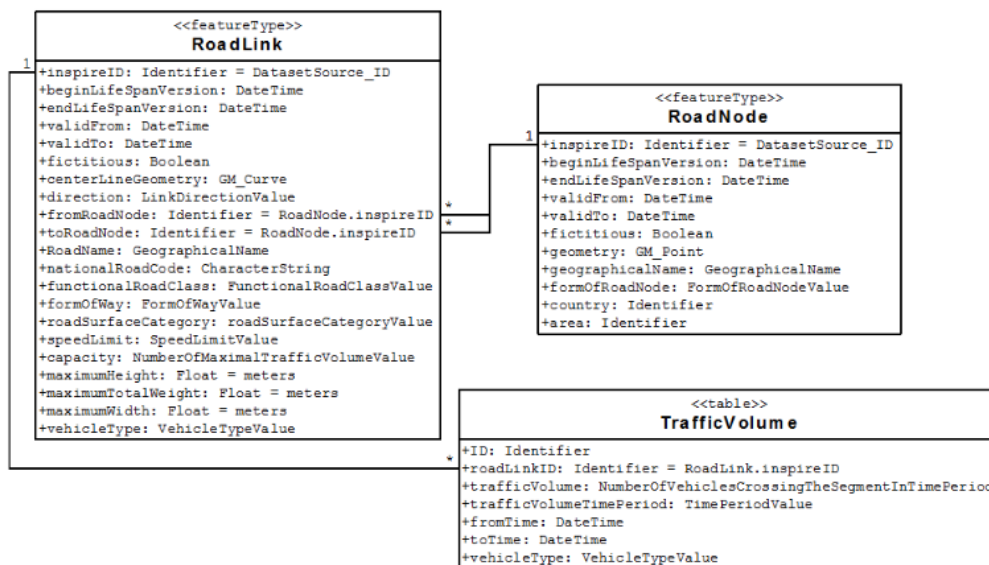
Tabulka 3: Ukázka OTM čísledníku.

RoadSurfaceCategoryValue	OSM.roads.surface
+ paved:	paved, asphalt, cobblestone, cobblestone:flattened, sett, concrete, concrete:lanes, concrete:plates, paving_stones, paving_stones:30, paving_stones:20, metal

RoadSurfaceCategoryValue	OSM.roads.surface
--------------------------	-------------------

+ unpaved:	<all other values>
------------	--------------------

Z technického hlediska byla původní data nejprve importována do RDMS PostgreSQL. Pomocí definovaného konverzního schématu byla vytvořena sada funkcí pro převod jednotlivých prvků do tzv. publikační databáze. Výsledná databáze slouží jako datový zdroj pro publikovaná data v příslušné formě pomocí softwaru jako UMN Mapserver, Geoserver případně dalšího.



Obrázek 10: Datový model OTM.

3.4.2 Převod dat do RDF

Převod dat OTM do podoby pětihvězdičkových LOD spočíval v několika krocích:

1. Určení identifikátorů.
2. Transformace dat do formátu RDF.
3. Doplnění vazeb na externí data a slovníky.

<u><<codeList>></u> <u>VehicleTypeValue</u>	<u><<codeList>></u> <u>TimePeriodValue</u>	<u><<codeList>></u> <u>LinkDirectionValue</u>	<u><<codeList>></u> <u>RoadSurfaceCategoryValue</u>
all vehicle bicycle car with trailer delivery truck emergency vehicle employee vehicle facility vehicle farm vehicle high occupancy vehicle light rail mail vehicle military vehicle moped motorcycle passenger car pedestrian private bus public bus residential vehicle school bus snow chain equipped vehicle tanker taxi transport truck trolley bus vehicle for disabled person vehicle with explosive load vehicle with other dangerous load vehicle with water polluting load	hour day monday, ..., sunday weekday weekend week month year	bothDirections inDirection inOppositeDirection	paved unpaved
	<u><<codeList>></u> <u>FormOfWayValue</u>	<u><<codeList>></u> <u>FormOfRoadNodeValue</u>	<u><<enumeration>></u> <u>FunctionalRoadClassValue</u>
	bicycleRoad dualCarriageway enclosedTrafficArea entranceOrExitCarPark entranceOrExitService freeway motorway pedestrianZone roundabout serviceRoad singleCarriageway slipRoad tractor trafficSquare walkway	enclosed traffic area junction level crossing pseudo node road end road service area roundabout traffic square	mainRoad firstClass secondClass thirdClass fourthClass fifthClass
		<u><<codeList>></u> <u>FormOfRoadNodeValue</u>	
		enclosed traffic area junction level crossing pseudo node road end road service area roundabout traffic square	

Obrázek 11: Číselníky použité v OTM.

Ad 1. Jako jednoznačný identifikátor bylo v případě jednotlivých komponent silnic a ulic převzato označení těchto komponent ve zdrojové sadě (OpenStreetMap). Pouze bylo upraveno do formy URI (což spočívalo v připojení jmenného prostoru `http://opentransportmap.info/rdf#`).

Ad 2. Data pro jednotlivé silniční úseky byla do RDF (viz následující ukázka kódu) transformována pomocí Bash skriptu, který používal jako vstupní soubor tabulková data ve formátu CSV. Vzhledem k neexistenci slovníků, obsahujících vhodné sémanticky popsané atributy, byl datový model vytvořen převážně transkripcí původních jmen atributů a přidáním jmenného prostoru `http://opentransportmap.info/rdf#` s prefixem `otm:`. Kromě tohoto jmenného prostoru, s jehož pomocí byly deklarovány atributy datového modelu, byly využity ještě slovníky GeoSPARQL (pro zápis geometrie ve formátu WKT) a Web Ontology Language (pro odkaz na identické prvky).

```
<rdf:Description
rdf:about="http://opentransportmap.info/rdf/10612539">
<geos:asWKT
rdf:datatype="http://www.openlinksw.com/schemas/virttrdf#Geometry">
LINESTRING (15.0848344 50.9267013,15.0839874 50.9265742)
</geos:asWKT>
<otm:beginLifeSpanVersion>2015/10/20</otm:beginLifeSpanVersion>
<otm:endLifeSpanVersion></otm:endLifeSpanVersion>
<otm:validFrom>2015/10/20</otm:validFrom>
<otm:validTo></otm:validTo>
<otm:fictitious>0</otm:fictitious>
<otm:direction rdf:resource="http://inspire.ec.europa.eu/codelist/
LinkDirectionValue/bothDirections"/>
<otm:fromRoadNode>6959932</otm:fromRoadNode>
<otm:toRoadNode>6959933</otm:toRoadNode>
<otm:functionalRoadClass rdf:resource="http://inspire.ec.europa.eu/
enumeration/FunctionalRoadClassValue/fifthClass"/>
<otm:formOfWay rdf:resource="http://inspire.ec.europa.eu/codelist/
FormOfWayValue/singleCarriageway"/>
<otm:roadSurfaceCategory rdf:resource="http://inspire.ec.europa.eu/
codelist/RoadSurfaceCategoryValue/unpaved"/>
<otm:capacity>2000</otm:capacity>
<otm:vehicleType rdf:resource="http://inspire.ec.europa.eu/
```

```
codelist/VehicleTypeValue/None"/>
<otm:z_order>0</otm:z_order>
<owl:sameAs
rdf:resource="http://linkedgedata.org/triplify/way29260269">
</rdf:Description>
```

Ad 3. Vazby na externí zdroje lze rozdělit do dvou skupin – vazby na externí data a vazby na slovníkové položky. V prvním případě byla pomocí vlastnosti `owl:sameAs` definována vazba na ekvivalentní prvek v datové sadě LinkedGeoData³, která obsahuje prvky OpenStreetMap převedené do podoby Linked Open Data. V druhém případě byly využity INSPIRE Registry⁴, které jsou k dispozici také ve formátu RDF. Proto byly příslušné položky seznamů přetřansformovány z textové podoby do formy URI odkazů do INSPIRE Registry.

Podobným způsobem jako cestní úseky byla do RDF transformována i data týkající se intenzity dopravy (příklad viz níže). Tato data nemají vlastní prostorovou složku, ale jsou připojena pomocí vlastnosti `otm:roadlink` k jednotlivým úsekům cest.

```
<rdf:Description
rdf:about="http://opentransportmap.info/rdf/Volume16618393">
<otm:roadlink
rdf:resource="http://opentransportmap.info/rdf/11836027"/>
<otm:trafficvolume>5</otm:trafficvolume>
<otm:column>333</otm:column>
<otm:from
mtime>2016-04-04 00:00:00</otm:fromtime>
<otm:totime>2016-04-04 01:00:00</otm:totime>
</rdf:Description>
```

³<http://linkedgedata.org>

⁴<http://inspire.ec.europa.eu/registry/>

3.5 Open Land Use Map

Open Land Use (OLU) Map neboli Otevřená mapa využití území je příkladem datové sady poskytující prostorová data využití území. OLU vznikla jako jeden z pilotních záměrů projektu SDI4Apps⁵. OLU je součástí iniciativy, která má za cíl vytvořit souvislou otevřenou mapu využití území pro celou Evropu v souladu s direktivou INSPIRE. V Evropě dosud neexistuje taková souvislá mapa, která je potřebná jak pro účely plánování, výzkumu, tak i pro potřeby investorů a obyvatel. Mapa vzniká na základě specifikací INSPIRE tématu Užití území a OLU České republiky (Obrázek 12) je díky dostupnosti otevřených podkladů jedním z prvních a nejlepších příkladů. Kromě ČR je v současné době pokryto Rakousko, Slovensko, Lotyšsko, Řecko, část Belgie (Flandry), Maďarsko, Španělsko.

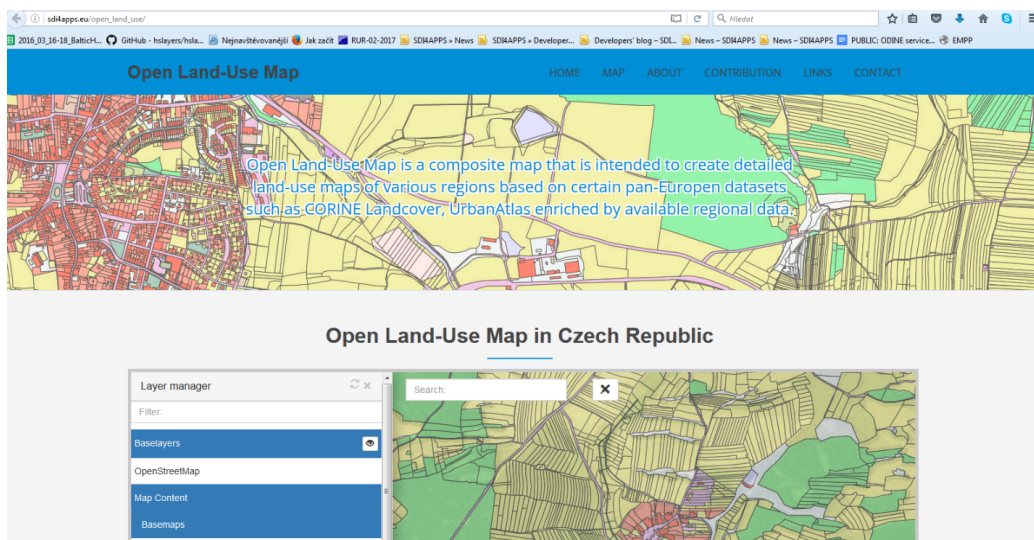
Pro účel projektu TB0500MV003 byl tento produkt vybrán z následujících důvodů: - Datový model je kompatibilní s INSPIRE - Jedná se o otevřená data - Jedná se o vektorová data - Část dat je zpřístupněna ve formátu RDF - Jedná se o rozsáhlý datový soubor - Data jsou dnes publikována dle několika úrovní hvězdičkové konvence

Datová sada vznikla kombinací dat digitální katastrální mapy, LPIS, Urban Atlas a CORINE land cover.

Z hlediska metodiky publikování otevřených dat se jedná o prostorovou datovou sadu, jež je k dispozici v několika různým podobách, které odpovídají různým úrovním pěti hvězdičkového modelu.

- As data (Obrázek 13) - V tomto případě je možné stáhnout jednotlivé části dat v proprietárním (avšak v geo-komunitě velice rozšířeném) formátu SHP. Při stažení je dostupné i permanentní URL příslušné datové sady. Tuto službu lze hodnotit jako dvouhvězdičkovou, protože jsou data strojově čitelná.
- As a map (Obrázek 14) - Zde jsou data dostupná v rastrové podobě pomocí formátu standardu WMS (označováno také jako prohlížečská služba). Výhodou je nabízená aplikace, která umožňuje přímé začlenění datové sady do konkrétní webové stránky. Tuto službu lze hodnotit jako jednohvězdičkovou, protože neumožňuje strojové čtení dat (jedná se

⁵<http://sdi4apps.eu/>



Obrázek 12: Open Land Use České republiky.

Download Open Land Use dataset in format of shapefile on level of municipalities.

Ceská Republika

Jihomoravský kraj

Břeclav

Diváky

Download

Obrázek 13: Open Land Use - as data.

de-facto o vektorová data, poskytovaná formou rastru). V případě jiného typu dat (např. ortofoto) by však tato metoda splňovala minimálně tři hvězdičky.



Obrázek 14: Open Land Use - as a map.

- As a service - datová sada ve formě WFS (stahovací služba) a ve formě Linked Open Dat (RDF). Tato úroveň zatím není na stránkách k dispozici. Metoda splňuje tři hvězdičky.
- As a RDF - OLU je transformováno do RDF schématu a publikováno prostřednictvím SPARQL endpoint.

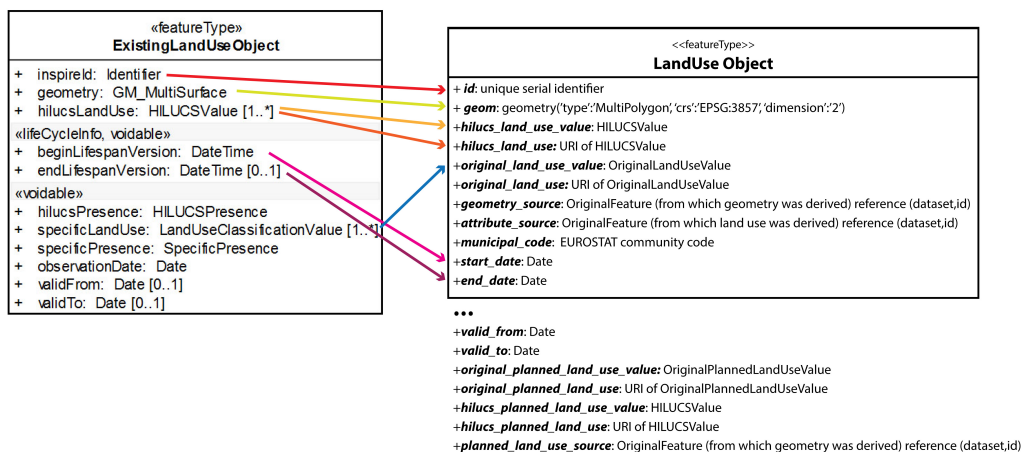
3.5.1 OLU datový model

Při tvorbě datového modelu se vycházelo z datové specifikace INSPIRE pro Land Use. Definice datového modelu byla zaměřena na současné užití území. Budoucí užití území není v této době v OLU modelu zahrnuto.

Model současného užití území byl oproti INSPIRE specifikaci rozšířen o několik prvků (Obrázek 15). Pro klasifikaci objektů byl užit kód HILUCS⁶.

Při implementaci OLU modelu jsou hlavními změnami přidání atributů `geometry_source` (zdroj geometrii prvku), `attribute_source` (zdroj podle kterého se určila třída využití území daného prvku) a `municipal_code` (kód

⁶<http://inspire.ec.europa.eu/codelist/HILUCSValue/>



Obrázek 15: OLU změny v modelu oproti INSPIRE.

obce dle číselníku Eurostatu). Atribut **geometry_source** má plnit funkci odkazu na prvek, z něhož byla odvozena geometrie prvku OLU.

Atribut **attribute_source** představuje odkaz na prvek, z něhož byla odvozena třída využití území daného prvku. Má stejný datový typ a strukturu jako atribut **geometry_source**. Důvodem je, že ne každý zdroj geometrii prvků obsahuje informaci o využití území.

Tyto dva atributy (**geometry_source** a **attribute_source**) jsou mimořádně důležité i z hlediska publikace dat. V tomto případě představují metadata o daných prvcích. V případě datových sad publikovaných státními orgány jako například jsou katastrální data - jejich metadata se publikují obvykle za celou datovou sadu, takže pak na hledání metadat se použije klíč **dataset**. V případě, že se jedná o datovou sadu podobnou otevřené databázi OpenStreetMap, jsou metadata na úrovni jednotlivých prvků.

Atribut **municipal_code** čili kód obci dle číselníku Eurostatu je užít pro potřebu jednoduššího filtrování a správy dat. Data jsou rozdělena podle obci NUTS 5.

V budoucnu, pokud se začnou publikovat územní plány měst v digitální formě, nebude problém rozšířit datový model o dodatečný atribut plánované využití území a související územně plánovací dokumentace.

Publikaci dat v RDF ilustruje následující příklad:

```

<?xml version="1.0" encoding="utf-8"?>
<rdf:RDF xmlns:geos="http://www.opengis.net/ont/geosparql#"
xmlns:olu="http://sdi4apps.eu/open_land_use/rdf#"
xmlns:elu=" http://inspire.ec.europa.eu/schemas/elu/4.0/"
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">
<rdf:Description rdf:about=
"http://sdi4apps.eu/open_land_use/rdf/1">
<geos:asWKT rdf:datatype="http://www.openlinksw.com/
schemas/virttrdf#Geometry">
POLYGON(1634627.98094257 6408270.48312739,1634612.40675096
6408267.16978162,1634586.00576481 6408261.19679501,
1634571.73370109 6408258.29300313,1634556.39552567
6408254.96435802,1634553.93945051 6408261.43751295,
1634548.23418306 6408276.46679834,1634556.17803332
6408278.17588873,1634556.08988778 6408278.60195522,
...
6408285.80885242,1634689.49425963 6408283.88115016,
1634671.13085769 6408279.84894462,1634643.04040244
6408273.68033685,1634636.39974587 6408272.25598002,
1634630.91114475 6408271.08031145,1634627.98094257
6408270.48312739)</geos:asWKT>
<elu:specificLandUse>ostatní komunikace</elu:specificLandUse>
<elu:hilucsLandUse rdf:resource="http://inspire.ec.europa.eu/
odelist/HILUCSValue/4_1_1_RoadTransport"/>
<olu:geometrySource
rdf:resource="http://vdp.cuzk.cz/vdp/ruian/parcely/2281304201"/>
<olu:attributeSource
rdf:resource="http://vdp.cuzk.cz/vdp/ruian/parcely/2281304201"/>
<olu:usedAttribute>způsob využití</elu:usedAttribute>
<elu:beginLifeSpanVersion>2016-09-28</elu:beginLifeSpanVersion>
<elu:endLifeSpanVersion></elu:endLifeSpanVersion>
<olu:municipalEurostatCode>CZ0201529303
</elu:municipalEurostatCode>
</rdf:Description>

```

3.6 Adresy

Problematika adres v Linked Open Data je klíčová, protože kromě nepřímého georeferencování konkrétních prostorových objektů mají adresy nezastupitelné místo také v metadatech. V rámci projektu SDI4Apps (pilotní aplikace Smart Points of Interest) byl proveden krátký průzkum zaměřený na standardy spojené se zápisem adres ve formátu RDF, které by mohly nahradit původní jednoduchou formu zápisu ve formě nestrukturovaného textového řetězce. Různé přístupy byly diskutovány na schůzce se zástupci ČÚZK i s účastníky projektů SDI4Apps, OTN a externími odborníky (obě schůzky se uskutečnily v lednu 2016).

Na základě získaných informací byl nakonec pro zápis adres ve SPOI zvolen slovník ISA Programme Location Vacabulary (Perego a Lutz 2015). Tento slovník je vhodný i pro ostatní prostorová data, přičemž jeho základní výhodou je možnost mapování do INSPIRE datového modelu pro adresy (součást INSPIRE specifikace Addresses⁷). Slovník obsahuje následující komponenty pro zápis adresy:

- **locn:fullAddress** - kompletní adresa zapsaná ve formě nestrukturovaného textového řetězce.
- **locn:poBox** - číslo poštovní přihrádky (P.O. BOX).
- **locn:thoroughfare** - označení pojmenovaného objektu (například ulice nebo třída).
- **locn:locatorDesignator** - v České republice se může jednat o číslo orientační, domovní, písmeno v orientačním čísle, typ domovního číslo, poštovní přihrádku
- **locn:locatorName** - například jméno budovy nebo celého komplexu.
- **locn:addressArea** - zpravidla část města, čtvrt nebo obec.
- **locn:postName** - jméno poštovního úřadu.
- **locn:adminUnitL2** - jméno kraje, federálního státu nebo jiného regionu.
- **locn:adminUnitL1** - zpravidla jméno státu, podle standardu je příkladem dobré praxe použití kódu podle ISO 3166-1.
- **locn:postCode** - poštovní směrovací číslo.
- **locn:addressId** - unikátní identifikátor podle datové specifikace směrnice INSPIRE.

⁷http://inspire.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_AD_v3.0.pdf

Příklady zápisu adresy pomocí ISA Programme Location Vocabulary lze najít v testovací aplikaci Core Location Pilot - Interconnecting Belgian address data⁸:

```
<rdf:RDF>
<rdf:Description rdf:about="http://location.testproject.eu/doc/
ad/AddressRepresentation/AGIV/2000017467">
<foaf:primaryTopic rdf:resource="http://location.testproject.eu/
so/ad/AddressRepresentation/AGIV/2000017467"/>
</rdf:Description>
<rdf:Description rdf:about="http://location.testproject.eu/id/
address/AGIV/2000017467">
<rdfs:seeAlso rdf:resource="http://location.testproject.eu/so/
ad/AddressRepresentation/AGIV/2000017467"/>
</rdf:Description>
<rdf:Description rdf:about="http://location.testproject.eu/so/
ad/AddressRepresentation/AGIV/2000017467">
<rdf:type rdf:resource="http://www.w3.org/ns/locn#Address"/>
<ns3:abstracts rdf:resource="http://location.testproject.eu/
id/address/AGIV/2000017467"/>
<ns4:adminUnit rdf:resource="http://location.testproject.eu/
so/au/AdministrativeUnit/STATBEL/23033"/>
<locn:adminUnitL4>ArrondissementHalle-Vilvoorde
</locn:adminUnitL4>
<locn:postName>Hoeilaart</locn:postName>
<ns6:transportLink rdf:resource="http://location.testproject.eu/
so/tn/Road/RN/15601625"/>
<locn:adminUnitL5>Hoeilaart</locn:adminUnitL5>
<locn:locatorDesignator>5</locn:locatorDesignator>
<locn:thoroughfare>Terhulpsesteenweg</locn:thoroughfare>
<ns3:osmVisualization>http://nominatim.openstreetmap.org/
search?q=Terhulpsesteenweg 5,Hoeilaart&polygon=1
</ns3:osmVisualization>
<locn:adminUnitL2>Vlaams Gewest</locn:adminUnitL2>
<locn:fullAddress>Terhulpsesteenweg 5 <br/> 1560 Hoeilaart
</locn:fullAddress>
```

⁸<http://location.testproject.eu/BEL/>

```

<locn:postCode
rdf:datatype="http://www.w3.org/2001/XMLSchema#integer">
1560</locn:postCode>
<ns6:addressFeature rdf:resource="http://location.testproject.eu/
so/ad/Address/AGIV/2000017467"/>
<locn:adminUnitL1
rdf:datatype="http://www.w3.org/2001/XMLSchema#integer">
0</locn:adminUnitL1>
<locn:adminUnitL3>Provincie Vlaams-Brabant</locn:adminUnitL3>
<locn:geometry rdf:resource="http://location.testproject.eu/id/
geometry/AGIV/2000017467"/>
</rdf:Description>
</rdf:RDF>

```

Kromě ISA Programme Location Vocabulary byly ještě diskutovány následující slovníky, ale jejich použití nebylo shledáno jako dostatečně efektivní a zároveň splňující všechny požadavky na prostorová data, včetně zohlednění směrnice INSPIRE.

- vCard⁹ - zřejmě světově nejpoužívanější slovník pro kódování adres, který je velmi jednoduchý (obsahuje pouze komponenty extended-address, locality, country-name, post-office-box, region, postal-code, street-address), avšak není možné bezproblémové propojení se směrnicí INSPIRE. Přístup založený na vCard používají například OxPoints¹⁰ nebo Organization ontology¹¹.
- GeoNames Ontology
- Postal Address Resource¹²
- FOAF¹³
- Addresses Ontology¹⁴

⁹<http://www.w3.org/Submission/vcard-rdf/>

¹⁰<http://help.it.ox.ac.uk/oxpoints/index>

¹¹<http://www.epimorphics.com/public/vocabulary/org.html>

¹²<https://www.oclc.org/developer/develop/web-services/worldcat-registry/rdf-interface/postal-address-resource.en.html>

¹³<http://xmlns.com/foaf/spec/>

¹⁴<http://www.omg.org/spec/EDMC-FIBO/FND/Places/Addresses.rdf>

Kapitola 4

Závěr

Metodika uvedená v tomto dokumentu byla vypracována v roce 2016 v rámci veřejné zakázky ve výzkumu, vývoji a inovacích TB0500MV003 s názvem Vypracování certifikované metodiky pro publikování prostorových informací ve formě otevřených dat. Podle zadání zakázky je cílem metodiky podpořit rozvoj podmínek pro efektivní využívání prostorových informací ve veřejných službách a službách veřejné správy a naplňovat principy otevřených dat veřejné správy i v této oblasti v návaznosti na systémové projekty k otevřeným datům v České republice. Na základě vypracované certifikované metodiky mohou následně být iniciovány potřebné legislativní změny souvisejících právních předpisů.

Metodika je určena především orgánům veřejné správy (některé příklady subjektů veřejné správy publikujících prostorová data jsou uvedeny v Příloze A), které se rozhodnou publikovat (zveřejňovat pro další nejen veřejnoprávní subjekty) prostorová data ve formě otevřených propojených dat (Linked Open Data). Veřejná správa však není jedinou skupinou, která může pocítovat výhody plynoucí z publikování otevřených prostorových dat, včetně návodných kroků, které mohou toto publikování usnadnit. Výstupy z metodiky mohou využívat nejen subjekty publikující data, ale také organizace data zpracovávající a dále využívající. Ty mohou využít například odkazy na slovníky a formáty, které mohou začlenit do svých služeb nebo aplikací a zajistit tak lepší sémantický popis nebo automatickou validaci. Dalším benefitem úzce souvisejícím s propojenými daty je možnost propojení dat veřejné správy s dalšími datovými zdroji, a tak získat nové odvozené informace.

Jak bylo poznamenáno v úvodní části metodiky, tento text je koncipován jako soubor technických kroků, které vedou k otevírání prostorových dat. Autoři neměli ambice obecně popisovat fenomén otevřených dat (v části Úvod je uvedeno velké množství zdrojů věnujících se otevřeným datům na obecné úrovni), ani argumentovat ve prospěch otevírání prostorových dat ve veřejné správě (výše uvedené dokumenty poskytují celou řadu argumentů podporujících otevírání dat, ale vlastní rozhodnutí je spíše na politické úrovni). Kromě toho, že se metodika věnuje prostorovým datům, její další omezení spočívá v orientaci na aktuální situaci v České republice.

Jednotlivé kroky vedoucí k otevírání prostorových dat vycházejí jednak z obecných standardů, příkladů dobré praxe a jednak z analýz současné situace v České republice i ve světě, včetně konzultací s vybranými subjekty statní správy (viz Příloha A). Vlastní metodika je členěna do pěti základních kapitol, které na jedné straně odpovídají jednotlivým segmentům pětihvězdičkového klasifikačního systému Linked Open Data (viz Úvod) a na straně druhé jsou zaměřeny na konkrétní problémy týkající se prostorových dat. Následující seznam ukazuje části metodiky a odpovídající úroveň pětihvězdičkového klasifikačního systému.

- Licence a ochrana osobních údajů (*)
- Metadata (*)
- Formáty prostorových dat (**, ***)
- Publikování datových modelů (**, ***)
- Identifikátory (****)
- Vazby na externí datové zdroje (*****)

Každý řešený aspekt otevřených prostorových dat je nahlížen ze tří pohledů:

1. Teorie
2. Doporučení
3. Ukázky dobré praxe (část Modelové aplikace)

Jak již bylo několikrát uvedeno, pro publikování otevřených dat existuje řada návodů a metodik. Následující výčet uvádí, v čem spočívá hlavní inovativní aspekt této metodiky:

1. Orientace na prostorová data, která jsou do jisté míry velmi specifická.
2. Metodika obsahuje konkrétní ukázky a příklady dobré praxe z projektů, na nichž se jednotliví autoři podíleli.

-
3. Doporučení na jedné straně zohledňují národní specifika, na straně druhé akcentují mezinárodní standardy a doporučení.
 4. Autoři si jsou vědomi, že propojená otevřená data (úroveň pěti hvězdíček) představují nejvyšší standard, ale pro mnoho typů otevřených prostorových dat je postačující publikování na úrovni dvou nebo tří hvězdíček (česká legislativa považuje za otevřená taková data, která jsou publikována pod otevřenou licencí v neproprietárním formátu; viz kapitola Slovník základních pojmů).

Pro publikování otevřených prostorových dat na úrovni dvou a tří hvězdíček lze použít veškeré obvyklé technologie v ČR. Zahrnuje to komerční platformy ESRI, Intergraph, AutoCad ale také další v současnosti velice populární open source nástroje, především pak MapServer a Geoserver. Složitější je publikace dat ve formátu RDF a Linked Open Data. K uchování RDF dat je potřeba použít RDF úložiště. K dotazování se nad úložišti, která uchovávají data v RDF formátu, je využíván jazyk SPARQL. Proto musí taková databáze pro přístup k RDF datům zajišťovat i funkci SPARQL endpointu, tj. podporovat dotazování v jazyku SPARQL. Jako příklad takového řešení je možno uvést například Virtuoso.

Požadavek na otevřená jako data označená třemi hvězdíčkami (tedy neproprietární data pod otevřenou licencí) je částečně v rozporu s praxí v České republice. Většina orgánů veřejné správy využívá komerční technologie (převážně produkty společnosti ESRI). Ty pracují formátem SHP, který je sice otevřený, ale také proprietární. Navíc je tento formát de facto standardem v oblasti geoinformačních technologií. Většina běžného programového vybavení dokáže s tímto formátem pracovat ve velké míře efektivněji než s nějakým neproprietárním formátem typu GML, který je otevřený, neproprietární a závazný z pohledu směrnice INSPIRE. Z tohoto důvodu by bylo vhodné definovat požadavek na otevřená data jako na data tzv. dva a půl hvězdíčková, přičemž klíčové je používání metadat, publikování pod otevřenou licencí, soulad s legislativou a využívání otevřeného datového formátu.

Na závěr je vhodné poděkovat Doc. Ing. Jiřímu Šímovi, CSc., a dalším anonymním přispěvatelům za kontrolu textu, připomínky a návrhy.

T A

Č R

„Tato metodika pro publikování prostorových informací ve formě otevřených dat byla vytvořena s finanční podporou TA ČR“.

Použitá literatura

„298/2016 Sb. Zákon, kterým se mění některé zákony v souvislosti s přijetím zákona o službách vytvářejících důvěru pro elektronické transakce, zákon č. 106/1999 Sb., o svobodném přístupu k informacím, ve znění pozdějších předpisů, a zákon č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon), ve znění pozdějších předpisů". 2016. Parlament České republiky.

Akční plán Strategie rozvoje infrastruktury pro prostorové informace v České republice do roku 2020. 2015. Ministerstvo vnitra, Český úřad zeměměřický a katastrální, Ministerstvo obrany, Ministerstvo životního prostředí, Ministerstvo pro místní rozvoj, Ministerstvo dopravy, Ministerstvo zemědělství, Ministerstvo financí.

Bauer, Florian, a Martin Kaltenböck. 2011. „Linked open data: The essentials". *Edition mono/monochrom, Vienna.*

Bechhofer, Sean, a Alistair Miles. 2009. „SKOS simple knowledge organization system reference". *W3C recommendation, W3C.*

Berners-Lee, Tim. 2006. „Linked data, 2006".

Brickley, Dan, a R Guha. 2014. „RDF Schema 1.1. W3C Recommendation (25 February 2014)". *World Wide Web Consortium.*

Chlapek, Dušan, Jakub Klímek, Jan Kučera, a Martin Nečaský. 2015. *Metodika publikace otevřených a propojitelných dat.* Technologická agentura ČR, Fond Otakara Motejla. http://opendata.vse.cz/tacr/mf/TD020277_Metodika_publicace_otevrenych_a_propojitelných_dat.pdf.

Čepický, Jáchym, Martin Landa, Radek Augustýn, Jan Cibulka, Jan Michálek, a Lucie Prunerová. 2015. *Otevírání geografických dat - Případová studie.*

-
- OpenGeoLabs s.r.o., IPR Praha. <http://opengeolabs.cz/otevrena-geodata/>.
- Čerba, Otakar, a Karel Jedlička. 2014. „Geomatic Concepts in Agriculture Thesauri". *Papers in Economics and Informatics*.
- . 2016. „Linked Forest". *Open Geosciences*.
- Čerba, Otakar, Karel Charvát, Tomáš Mildorf, Raitis Bērziņš, Pavel Vlach, a Barbora Musilová. 2016. „SDI4Apps Points of Interest Knowledge Base". In *Progress in Cartography*, 229–37. Springer.
- ČSN P ISO/TS 19104 (979823) *Geografická informace - Terminologie*. 2010. Úřad pro technickou normalizaci, metrologii a státní zkušebnictví.
- Dean, Mike, Guus Schreiber, Sean Bechhofer, Frank van Harmelen, Jim Hendler, Ian Horrocks, Deborah L McGuinness, Peter F Patel-Schneider, a L Andrea Stein. 2004. „OWL web ontology language reference". *W3C Recommendation February 10*.
- Hakrae, Kim. 2012. „Overview of Open Data, Linked Data and Web Science".
- INSPIRE. 2010. „Technical Guidance to implement INSPIRE Discovery Services, version 2.12".
- INSPIRE, INS MD IMPL. 2016. „INSPIRE Metadata Implementing Rules: Technical Guidelines based on ISO/TS 19139:2007, v2.0".
- INSPIRE, INSMD. 2008. „Commission Regulation (EC) No 1205/2008 of 3 December 2008 implementing Directive 2007/2/EC of the European Parliament and of the Council as regards metadata (Text with EEA relevance)". *See also Corrigendum to INSPIRE Metadata Regulation*.
- INSPIRE, INSNS. 2009. „Commission Regulation (EC) No 976/2009 of 19 October 2009 implementing Directive 2007/2". *EC of the European Parliament and of the Council as regards the Network Services*.
- Kavouras, Marinos, a Margarita Kokla. 2007. *Theories of geographic concepts: ontological approaches to semantic integration*. CRC Press.
- Kovacs, K, C Dolbear, a J Goodwin. 2007. „Spatial concepts and OWL issues in a topographic ontology framework". *Proc. of the GIS*. Citeseer.
- Kritikos, Kyriakos, Yannis Rousakis, a Dimitris Kotzinos. 2013. „Linked open GeoData management in the cloud". In *Proceedings of the 2nd international*

workshop on open data, 3. ACM.

Kuhn, Werner, Tomi Kauppinen, a Krzysztof Janowicz. 2014. „Linked data-A paradigm shift for geographic information science". In *International Conference on Geographic Information Science*, 173–86. Springer.

Lebo, Timothy, a Gregory Todd Williams. 2010. „Converting governmental datasets into linked data". In *Proceedings of the 6th International Conference on Semantic Systems*, 38. ACM.

Longley, P. A., M. F. Goodchild, D. J. Maguire, a D. W. Rhind. 2001. *Geographic information systems and science*. John Wiley & Sons Ltd.

Metadata, Technická pracovní skupina. 2015. „Metadatový profil ČR pro soubory prostorových dat, sérií souborů prostorových dat a služeb založených na prostorových datech, verze 3.0".

Miles, Alistair, a Sean Bechhofer. 2009. „SKOS simple knowledge organization system reference". *W3C recommendation* 18: W3C.

Mráček, Jakub, a kolektiv. 2014. *Jak otevírat data*. Fond Otakara Motejla. <http://www.otevrenadata.cz/res/data/001/003498.pdf>.

Open Data Consultancy. Final Report. 2013. Scottish Government by APS Group Scotland.

Perego, Andrea, a Michael Lutz. 2015. *ISA Programme Location Core Vocabulary*. EU ISA Programme Core Vocabularies Working Group (Location Task Force). <https://www.w3.org/ns/locn>.

Perry, M., a J. Herring. 2012. „OGC GeoSPARQL-A geographic query language for RDF data". *OGC Implementation Standard*.

Pražák, Josef, a ostatní. 2005–2016. *Terminologický slovník zeměměřictví a katastru*. Vyzkumny ústav geodetický, topografický a kartografický, Terminologická komise Českého úřadu zeměměřického a katastrálního.

Raamkumar, Aravind Sesagiri, Muthu Kumaar Thangavelu, a others. 2015. „Designing a Linked Data Migrational Framework for Singapore Government Datasets". *arXiv preprint arXiv:1504.01987*.

Shadbolt, Nigel, Kieron O'Hara, Tim Berners-Lee, Nicholas Gibbins, Hugh Glaser, Wendy Hall, a others. 2012. „Linked open government data: Lessons

from data. gov. uk". *IEEE Intelligent Systems* 27 (3): 16–24.

Šíma, Jiří. 2016. *Terminologický výkladový slovník pro potřeby realizace Akčního plánu Strategie rozvoje infrastruktury pro prostorové informace v České republice do roku 2020 – verze 3.4.*

W3C. 2012. „OWL 2 web ontology language document overview". Word Wide Web Consortium.