

TAČR TB0500MV003: Vypracování certifikované metodiky pro publikování prostorových informací ve formě otevřených dat

Otakar Čerba (ZČU), Štěpán Kafka (HSRS), **Karel Charvát** (HSRS), Jan Ježek (ZČU), Stanislav Holý (HSRS), Václav Čada (ZČU), Dmitrii Kožuch (HSRS), Tomáš Mildorf (ZČU)

Seminář GeolInfoStrategie a TA ČR - Spolupráce Technologické agentury ČR při implementaci GeolInfoStrategie, 26. dubna 2017, Praha

Cíle projektu

- ▶ Vytvořit **certifikovanou metodiku pro publikování prostorových informací ve formě otevřených dat** s cílem podpořit rozvoj podmínek pro efektivní využívání prostorových informací ve veřejných službách a službách veřejné správy.
- ▶ Naplňovat principy otevřených dat veřejné správy i v této oblasti v návaznosti na systémové projekty k otevřeným datům v ČR.

Výstupy projektu

1. Analýza a definice uživatelských požadavků (především veřejná správa, ale i občané a komerční sféra; rozhovory se 4 klíčovými organizacemi)
 - ▶ Institut plánování a rozvoje Hlavního města Prahy
 - ▶ Český úřad zeměměřický a katastrální
 - ▶ Český statistický úřad
 - ▶ Agentura ochrany přírody a krajiny
2. Analýza stávající situace (existující systémy v ČR i zahraničí, analýza legislativní, technické, sémantické a organizační stránky publikace otevřených prostorových dat)
3. Postupy a opatření pro publikaci prostorových dat jako dat otevřených (metodika)
4. Modelové ukázky publikace prostorových dat jako dat otevřených

Uživatelé

- ▶ **Organizace veřejné správy** publikující prostorová data jako data otevřená nebo jako Linked Open Data
- ▶ Subjekty mimo veřejnou správu, které pracují s otevřenými prostorovými daty (vývoj aplikací nebo služeb, propojování a vizualizace dat) nebo taková data vytváří a publikují

Osnova výstupního dokumentu (metodiky)

- ▶ Seznam zkratk – více než 100 zkratk souvisejících s problematikou otevřených a prostorových dat
- ▶ Úvod – představení problematiky otevřených dat a jejich vztahu k prostorovým datům
- ▶ Slovník základních pojmů – vymezení základních pojmů (například termíny otevřená a prostorová data) a možnost porovnání různých způsobů definic
- ▶ Návrhová část metodiky – přehled klíčových kroků pro publikování otevřených dat s ohledem na specifika prostorových dat
- ▶ Modelové aplikace
- ▶ Závěr
- ▶ Příloha A. Analytická část metodiky

Jak s metodikou pracovat

- ▶ Nejedná se o souvislý text
- ▶ Jednotlivé kapitoly se zabývají klíčovými problémy při publikování otevřených dat s ohledem na data prostorová
- ▶ Uživatel může využívat pouze ty části, které považuje za vhodné, aby dosáhl konkrétního stupně otevřenosti a propojenosti dat
- ▶ Každá kapitola má stejnou strukturu – teoretická pasáž (popis problému), doporučení a ilustrativní příklad v části Modelové aplikace

Jak byly voleny kapitoly metodiky

- ▶ Jednotlivé kapitoly se zabývají klíčovými komponentami otevřených dat s ohledem na tzv. pětihvězdičkový systém hodnocení propojených otevřených dat:
 - ▶ Publikování pod otevřenou licenci (*)
 - ▶ Publikování ve strojově čitelném formátu (**)
 - ▶ Publikování pomocí otevřeného, neproprietárního formátu (***)
 - ▶ Publikování ve formátu RDF (****)
 - ▶ Vytvoření linků na externí datové zdroje (*****)

Kapitoly návrhové části metodiky

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

Licence a ochrana osobních údajů

- ▶ Ochrana osobních údajů - stručný rozbor zákona č. 101/2001 Sb., o ochraně osobních údajů
- ▶ Práva duševního vlastnictví a otevřené licence – Sdělení Evropské komise 2014/C 240/01 doporučuje používat normované mezinárodní otevřené licence **Creative Commons** – CC0 Public Domain, Creative Commons s omezeními CC-BY (attribution) a CC-BY-SA (attribution and share-alike)

Metadata

- ▶ Požadavky na metadata
 - ▶ INSPIRE, GeoDCAT
 - ▶ Národní katalog INSPIRE, Národní katalog otevřených dat (scénáře publikování)
- ▶ Popis řešení jednotlivých částí metadat (například klíčová slova, licence, adresy. . .)
- ▶ Ukázky v ISO 19139 a GeoDCAT

Formáty prostorových dat

- ▶ Přehled používaných formátů a software
- ▶ Hlavní otevřené datové formáty pro prostorová data
 - ▶ Vektorové – KML, GML, SHP, (Geo)JSON
 - ▶ Rastrové
 - ▶ Webové služby

Publikování datových modelů

- ▶ Složitost datových modelů
 - ▶ Jednoduchý datový model
 - ▶ Středně složitý datový model
 - ▶ Komplexní datový model
- ▶ Prostředky pro popis datových modelů - validovatelné (XML schémata) x nevalidovatelné (UML)

Identifikátory

- ▶ Metodika doporučuje stanovit systém URI pro základní registry (návrhy pro Registr osob, Registr obyvatel, Registr územní identifikace, adres a nemovitostí, Registr práv a povinností)
 - ▶ Dereferencovatelnost (vazba na RDF data)
 - ▶ V gesci správců registrů
 - ▶ Pod doménou data.gov.cz

Vazby na externí datové zdroje

- ▶ Základ Linked Data - sociální funkce
- ▶ Typy vazeb
 - ▶ Identické (SKOS, OWL, RDFS)
 - ▶ Topologické (GeoSPARQL - Region Connection Calculus 8)
 - ▶ Hierarchické a mereologické
- ▶ Hodnocení vhodnosti jednotlivých standardů - rozhodování podle vágnosti definice vazeb
- ▶ Ukázky - SPOI, OTM

Modelové aplikace

- ▶ Existující nebo v současnosti vyvíjené aplikace, slovníky nebo datové sady
 - ▶ Metadata (online převodník GeoDCAT a INSPIRE)
 - ▶ Ontologie VF XML (vazba na zdroje s explicitní sémantikou)
 - ▶ Smart Points Of Interest + SPOI Ontology (vazba na standardy a slovníky, vrstva s bodovou geometrií)
 - ▶ Open Transport Map (vazba na INSPIRE, vrstva s liniovou geometrií)
 - ▶ Open Land Use Map (vazba na INSPIRE, vrstva s areálovou geometrií)
 - ▶ adresy (vazba na standardy W3C, nepřímé georeferencování)
- ▶ Návaznost na evropské projekty (SDI4Apps, OTN)
- ▶ Ověření některých částí metodiky (vazba na INSPIRE, transformace metadat, efektivita zápisu prostorových dat v RDF...)

Závěr


- ▶ Metodika je koncipována jako soubor technických kroků, které vedou k otevírání prostorových dat v ČR
- ▶ Text neslouží k
 - ▶ vysvětlení obecné problematiky otevřených dat
 - ▶ argumentaci ve prospěch otevřených dat
- ▶ Pro publikování otevřených prostorových dat na úrovni dvou a tří hvězdiček lze použít veškeré obvyklé technologie v ČR
- ▶ Požadavek na otevřená jako data označená třemi hvězdičkami (neproprietární data pod otevřenou licencí) je částečně v rozporu s praxí v ČR (SHP)
- ▶ **Otevřená prostorová data v ČR jako na data tzv. dva a půl hvězdičková, přičemž klíčové je používání metadat, publikování pod otevřenou licencí, soulad s legislativou a využívání otevřených datových formátů**

Děkujeme za pozornost

HS-RS

 **geomatika**

 **ZÁPADOČESKÁ
UNIVERZITA
V PLZNI**

 **KATEDRA
APLIKOVANÉ
GEOMATIKY A GIS
ZÁPADOČESKÉ
UNIVERZITY
V PLZNI**

Program **Beta**

**T A
Č R**

Projekt TAČR TB0500MV003 je řešen s finanční podporou TA ČR