

SMART POINTS OF INTEREST

Otakar ČERBA¹, Michal KEPKA², Jan MACURA³, Jáchym KELLAR⁴

¹ Katedra geomatiky, Fakulta aplikovaných věd, Západočeská univerzita, Univerzitní 8, 306 14 Plzeň, Česká republika
cerba@kgm.zcu.cz

² Katedra geomatiky, Fakulta aplikovaných věd, Západočeská univerzita, Univerzitní 8, 306 14 Plzeň, Česká republika
mkepka@kgm.zcu.cz

³ Katedra geomatiky, Fakulta aplikovaných věd, Západočeská univerzita, Univerzitní 8, 306 14 Plzeň, Česká republika
macurajan@gmail.com

⁴ Katedra geomatiky, Fakulta aplikovaných věd, Západočeská univerzita, Univerzitní 8, 306 14 Plzeň, Česká republika
kellar@students.zcu.cz

Abstrakt

Smart Points of Interest (SPOI) je datová sada obsahující body zájmu. SPOI data jsou šířena ve formě otevřených propojených dat (Linked Open Data). To podporuje vývoj různých aplikací a služeb využívající SPOI. Tento příspěvek představuje nejen samotnou datovou sadu a její potenciál, ale také ukazuje aplikace postavené nad SPOI daty. jedná se například o mapový klient, SPARQL endpoint, aplikaci pro sběr dat, vyhledávání dat nebo generování webových stránek využívající SPOI.

Abstract

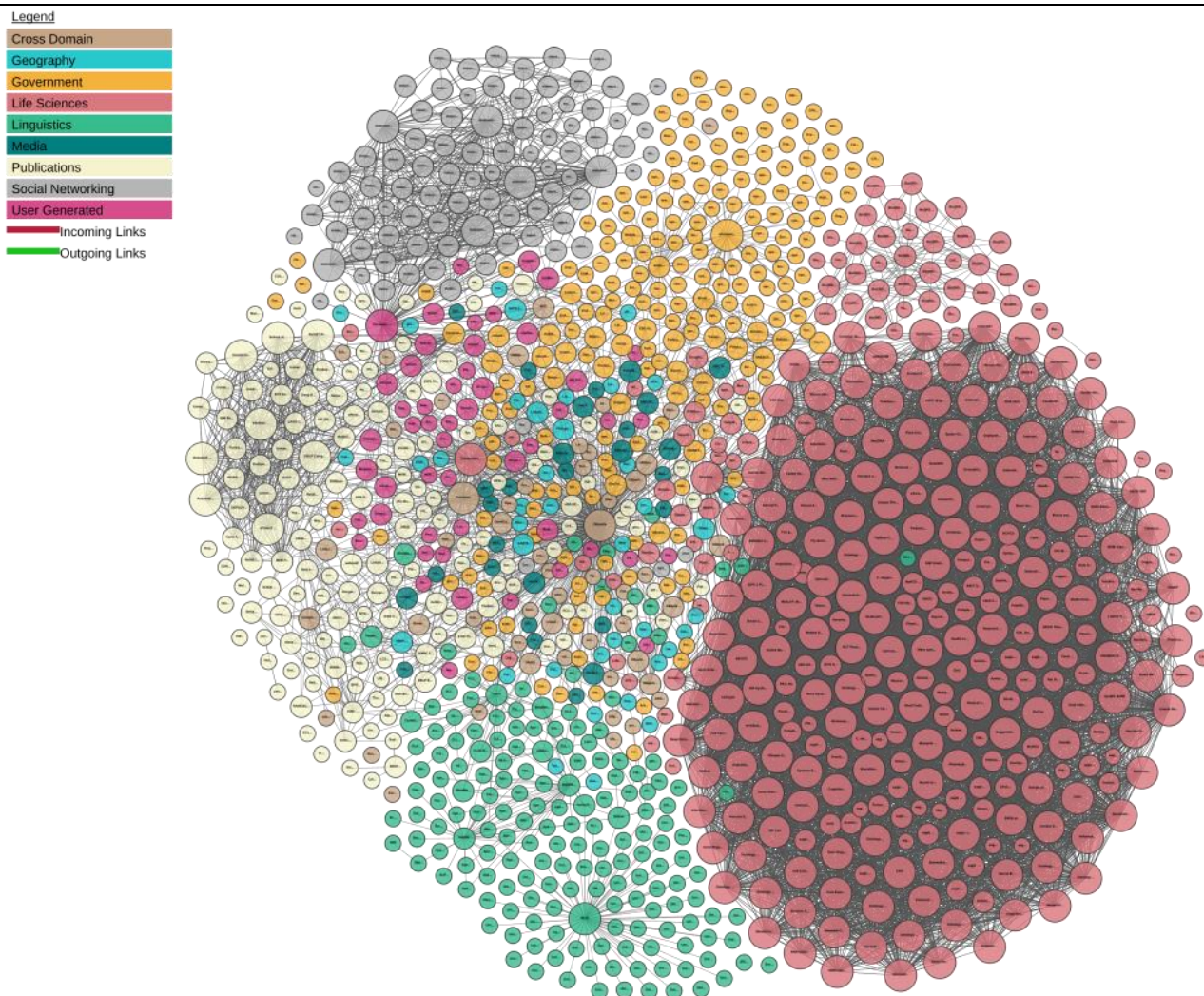
Smart Points of Interest (SPOI) is data set containing "points of interest". SPOI data is distributed as Linked Open Data. This approach support development of various applications and services using SPOI. This presentation introduces not only the SPOI data and its capacity, but also examples of applications built upon SPOI. They include for example map client, SPARQL endpoint, application for data collecting, data searching or generating of web pages using SPOI.

Klíčová slova: propojená data; prostorová data; SPOI; webová aplikace.

Keywords: Linked data; spatial data; SPOI; web application.

1 ÚVOD

Přístup Linked data (v českém jazyce se začíná prosazovat termín propojená data) vznikl před více než deseti lety (Berners-Lee, 2006). Jejich princip spočívá ve využívání technologií, standardů a postupů známých z prostředí webu pro publikování strukturovaných dat, nikoli pouze webových stránek ve formě víceméně volného textu (více o technologickém pozadí Linked data v publikacích Berners-Lee, 2006, Bizer et al., 2009 nebo Heath & Bizer, 2011). Ani přes poměrně důležitý podíl prostorových dat v Linking Open Data cloud diagramu (tyrkysově podbarvené položky v obrázku 1) není možné najít velké množství skutečných implementací propojených dat v oblasti geoinformačních technologií. A to i přes nesporné výhody Linked data, kterými jsou například zavedení jednoznačných identifikátorů pro jednotlivé objekty prostorových dat, požadavek na jasné deklarování licence, nutnost publikování dat na internetu, využívání neproprietárních formátů, doplnění explicitní sémantiky nebo propojení na externí datové zdroje.



Obr. 1. Linking Open Data cloud diagram (lod-cloud.net¹)

Článek Smart Points of Interest je rozdělený na dvě základní části. Nejprve je stručně popsána datová sada Smart Points of Interest (SPOI), která představuje jednu z celosvětově největších otevřených datových sad využívající Linked data. Data SPOI, která jsou vyvíjena v rámci projektu SDI4Apps především na Katedře geomatiky Západočeské univerzity v Plzni. SPOI obsahují zhruba 27 miliónů zájmových bodů, popsanych pomocí datového modelu využívajícího standardy organizací World Wide Web Consortium, Open Geospatial Consortium a dalších. Data jsou v současné době publikována ve formě mapové aplikace (pro prohlížení dat) i ve formě tzv. SPARQL endpoint, který představuje „stahovací službu“ fungující na principu dotazů, jejichž výstupem jsou prostorová data ve formátech CSV (Comma-separated values), RDF (Resource Description Framework) nebo JSON (JavaScript Object Notation). Druhou částí příspěvku je představení několika aplikací (v různých fázích vývoje), které se datovou sadou SPOI dokáží pracovat a využívat její přednosti.

2 SMART POINTS OF INTEREST

1 Linking Open Data cloud diagram 2017, by Andrejs Abele, John P. McCrae, Paul Buitelaar, Anja Jentzsch and Richard Cyganiak. <http://lod-cloud.net/>

Jak bylo uvedeno v úvodu, v rámci projektu SDI4apps – Uptake Of Open Geographic Information Through Innovative Services Based On Linked Data (sdi4apps.eu) 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 (point of interest), které jsou určeny jako datový podklad primárně pro aplikace zabývající se turistikou nebo cestovním ruchem. Informace uložené ve SPOI mohou sloužit i jiných oblastech, například jako velice účinná a efektivní propagace konkrétních památek nebo služeb.

Data SPOI nejsou uložena v tradiční relační databázi, ale jako grafová strukturu ve formě RDF (Resource Description Framework) trojice (viz ukázka kódu). Tento formát je v souladu s požadavky na tzv. pětihvězdičková Linked data, kde předpokladem k získání čtvrté hvězdy je identifikace datových položek pomocí URI (Uniform Resource Identifier) a zápis dat v univerzálním RDF formátu.

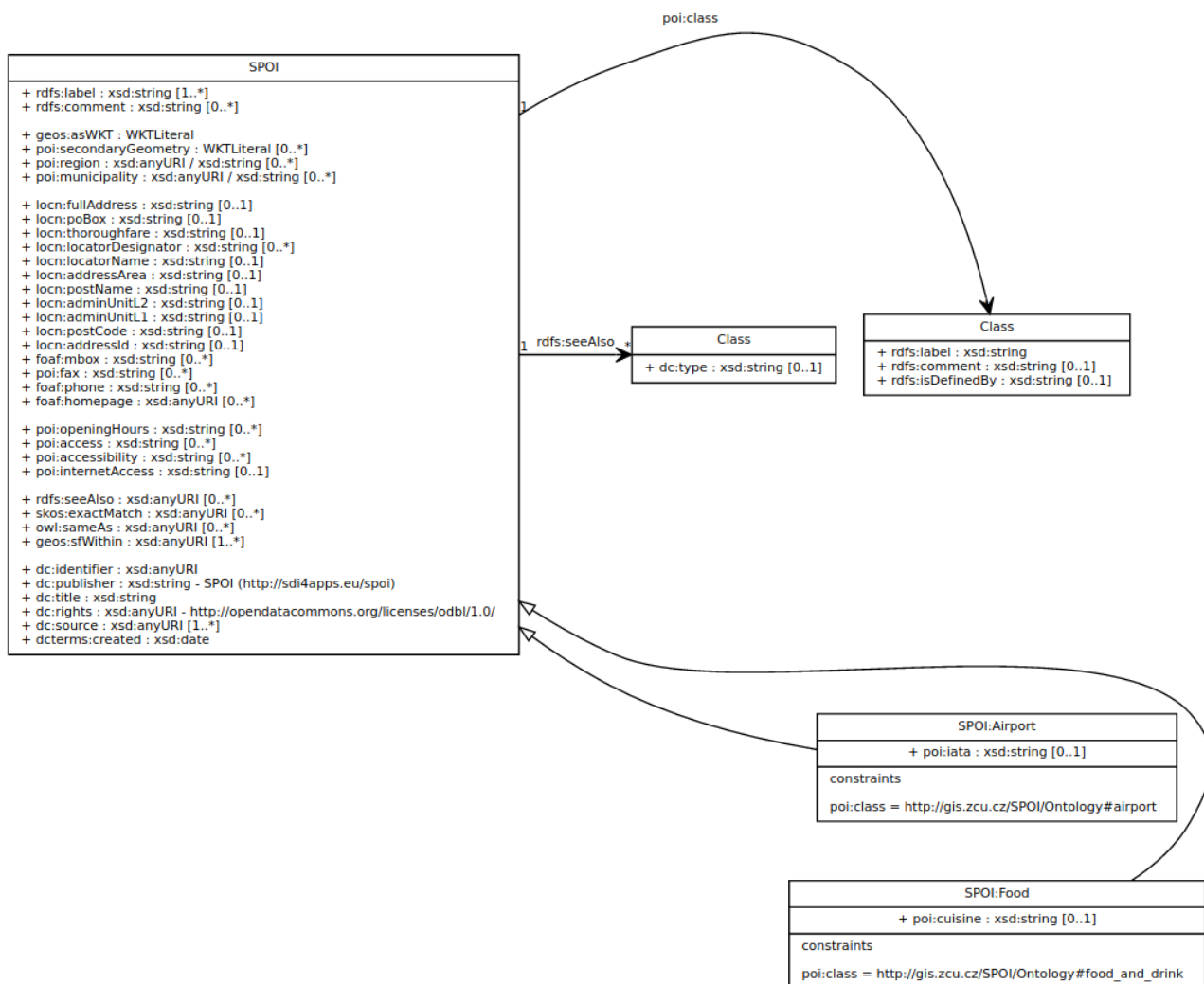
```
<rdf:Description rdf:about="http://www.sdi4apps.eu/poi/#COTM_+1+_1">
<rdfs:label>Škoda Park Doudlevce</rdfs:label>
<poi:class rdf:resource="http://gis.zcu.cz/SPOI/Ontology#climbing_wall"/>
<poi:class rdf:resource="http://gis.zcu.cz/SPOI/Ontology#culture_and_entertainment"/>
<geos:asWKT
rdf:datatype="http://www.openlinksw.com/schemas/virttrdf#Geometry">POINT(13.390697
49.722826)</geos:asWKT>
<geos:sfWithin rdf:resource="http://dbpedia.org/resource/Czech_Republic"/>
<geos:sfWithin rdf:resource="http://www.geonames.org/3077311"/>
<dc:identifier rdf:resource="http://www.sdi4apps.eu/poi/#COTM_+1+_1"/>
<dc:publisher>SPOI (http://sdi4apps.eu/spoi)</dc:publisher>
<dc:title>Škoda Park Doudlevce</dc:title>
<dc:rights rdf:resource="http://opendatacommons.org/licenses/odbl/1.0"/>
<dc:source rdf:resource="http://www.citadelonthemove.eu"/>
<dcterms:created rdf:datatype="http://www.w3.org/2001/XMLSchema#date">2016-11-
14</dcterms:created>
</rdf:Description>
```

V současné době je k dispozici více než 27 miliónu 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 (obě aplikace jsou popsány níže). 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,
- 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, Natural Earth nebo GeoNames.org. Kromě těchto globálních datových sad jsme zpracovali také lokální datové sady od partnerů projektů Citadel on the Move (cca 30 lokálních datových sad), SDI4apps (Zemgale, Pošumaví), FOODIE (Belluno) a OpenTransportNet (Antverpy, Issy).

Velkou výhodou řešení SPOI oproti konkurenčním datovým sadám je kromě faktu, že se jedná o otevřená a bežešvá data, především existence jednotného datového modelu (Obrázek 2). Ten maximálně využívá existující slovníky Linked Data a umožňuje snadné propojení s dalšími objekty.



Obr. 2. Datový model SPOI (verze z února 2017).

Veškeré informace o datech SPOI jsou k dispozici na webové stránce gis.zcu.cz/spoi. Zájemci zde mohou nalézt nejen odkazy na vizualizaci dat a možnost stažení dat, ale také aktuální verzi datového modelu, schéma datové harmonizace, seznam realizovaných změn, prezentaci představující SPOI a metadata ve formě DOAP (Description of a Project) a VoID (Vocabulary of Interlinked Datasets). Další informace o SPOI, především týkající se harmonizace a integrace dat z různých zdrojů a popisu tvorby datového modelu jsou k dispozici například v publikacích Čerba et al., 2016a nebo Čerba et al., 2016b.

3 VÝVOJ APLIKACÍ VYUŽÍVAJÍCÍCH SPOI

3.1 SPARQL endpoint

SPOI data jsou uložena v prostředí Virtuoso (virtuoso.openlinksw.com), které představuje moderní (tzv. hybridní) serverové řešení umožňující kromě tradičních relačních databází ukládat také grafové struktury ve formátu RDF. Z tohoto důvodu je tento nástroj populární ve velkém množství datových sad využívajících přístup Linked data (například DBpedia nebo LinkedGeoData).

Pro uživatele SPOI není ani zdaleka tak důležitý způsob uložení dat, ale především jejich dostupnost. Ta je zajištěna pomocí tzv. SPARQL endpoint. Jedná se o klienta dotazovacího jazyka SPARQL (Harris et al., 2013), který na základě uživatelského dotazu vygeneruje požadovaná data. Dotazovací jazyk SPARQL je podobný tradičnímu SQL (Structured Query Language) nejen posledními písmeny zkratky, ale také strukturou dotazu. Následující ukázka představuje jednoduchý dotaz vypíše identifikátory (ve formě URI) pro prvních 100 bodů zájmu v databázi SPOI.

```
PREFIX dc: <http://purl.org/dc/elements/1.1/>

select ?SPOI

where {?SPOI dc:source ?category}

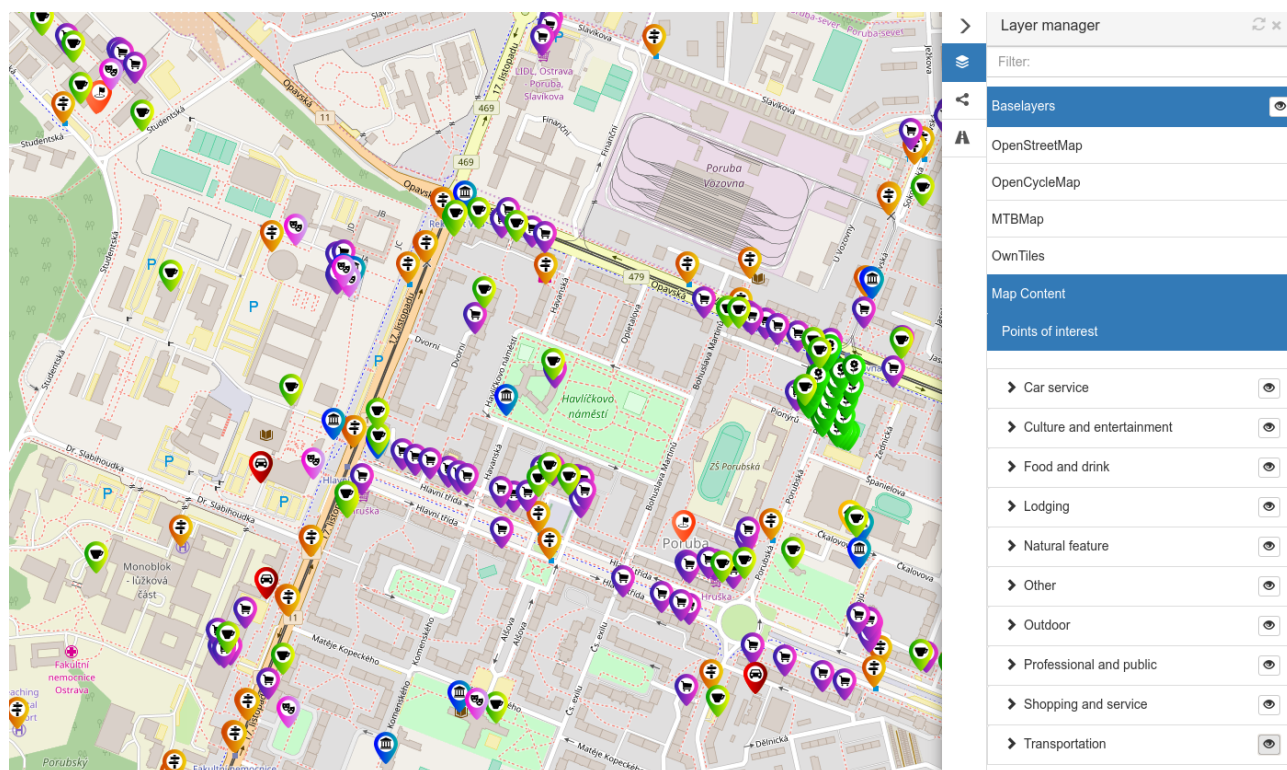
LIMIT 100
```

SPARQL endpoint poskytuje data v mnoha formátech jako například CSV, různé syntaxe RDF, JSON, HTML (HyperText Markup Language), XML (Extensible Markup Language), TSV (Tab-Separated Values) a podobně. Vzhledem k tomu, že se jedná o strukturovaná a v mnoha případech i kvalitně sémanticky popsaná data, není problém jejich transformace do některého z nebinárních formátů (například KML – Keyhole Markup Language nebo GML – Geography Markup Language) určených pro geografické informační systémy. Navíc většina současných GIS nástrojů dokáže zpracovat i formát CSV.

3.2 Mapový klient

Podobně jako SPARQL endpoint je z webové stránky SPOI dostupný i mapový klient (Obrázek 3). Klient je založený na technologii HS Layers, která představuje prostředí pro tvorbu mapových aplikací využívající především javascriptové knihovny OpenLayers a ExtJS, a je šířena jako open source pod GNU/GPL licenci.

SPOI klient není pouze tradičním vizualizačním nástrojem, který umožňuje pohyb v mapě, změnu měřítka nebo zapínání jednotlivých vrstev. Uživatel má možnost měnit grafický vzhled jednotlivých zobrazovaných dat a také data editovat či přidávat nové body. Veškeré změny jsou prozatím uložené v nové datové sadě a do SPOI budou exportovány během pravidelných aktualizací databáze, která probíhá čtyřikrát ročně.

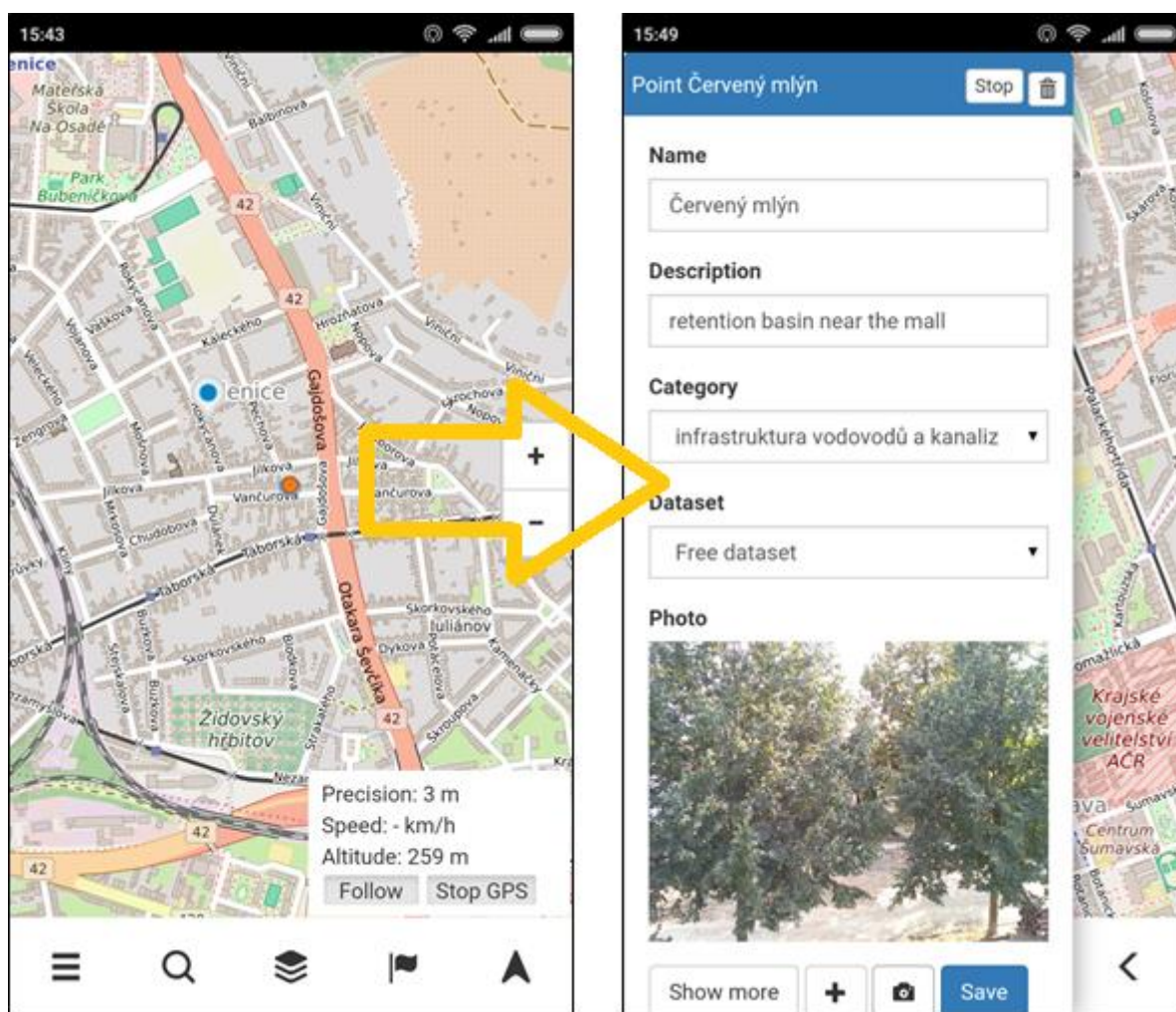


Obr. 3. Mapový klient SPOI (data z prosince 2016).

3.3 HSLCollector – aplikace pro sběr dat

Mobilní aplikace HSLCollector (Obrázek 4) v kombinaci se serverovou aplikací SensLog poskytuje dostupné řešení pro sběr a aktualizaci VGI (Volunteered Geographic Information). Pilotní testování tohoto řešení probíhá nad datovou sadou SPOI v rámci projektu SDI4Apps. Navržené řešení umožňuje přímo v terénu sbírat bodové jevy s různými vlastnostmi a ukládat je v serverovém úložišti.

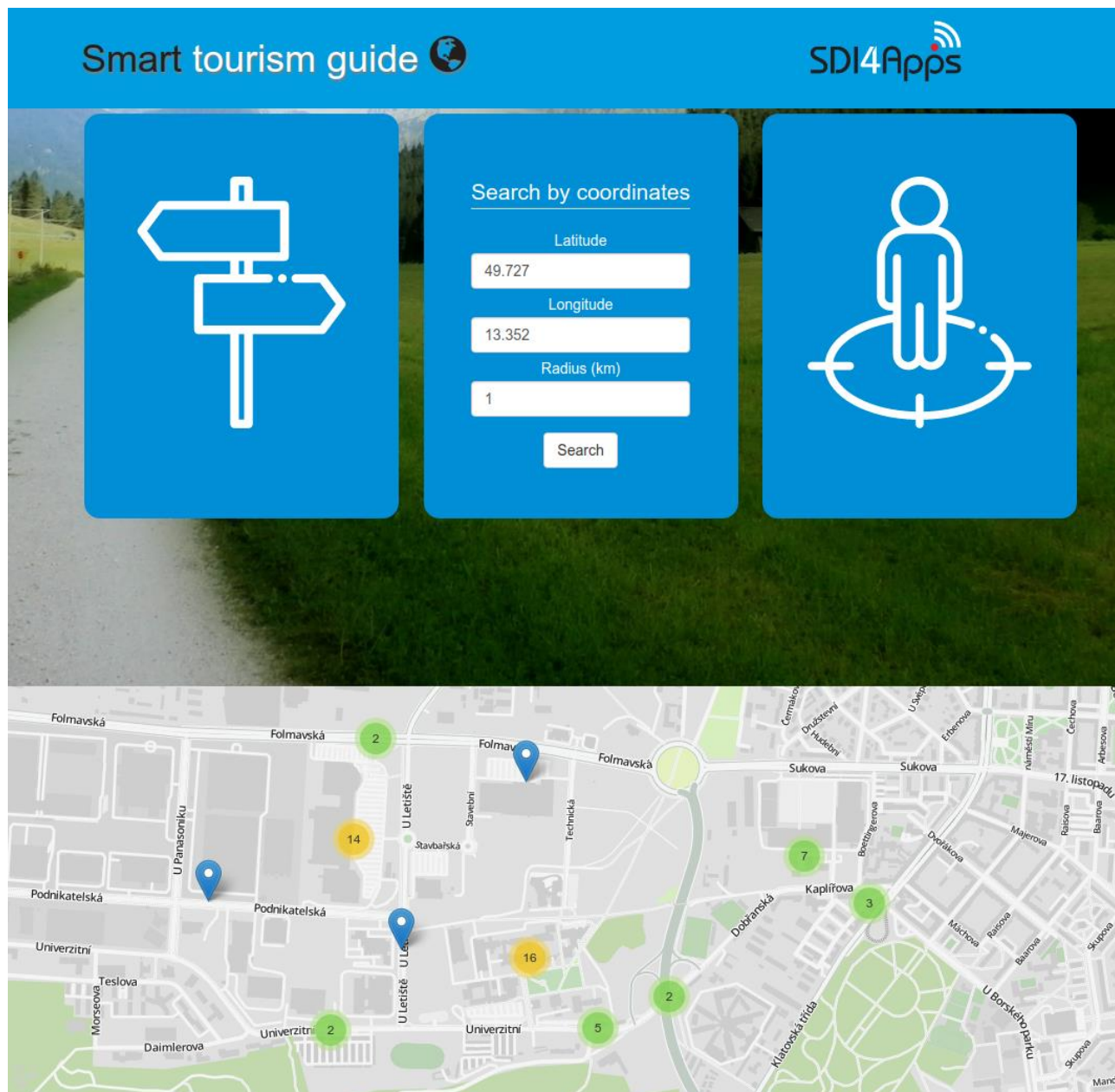
Každý jev je reprezentován bodem s prostorovou polohou určenou souřadnicemi a libovolným počtem atributů. Atributy jevu mohou být zaznamenány pomocí primitivních datových typů (číselné, textové apod.) nebo pomocí multimediálních dat (obrazová, zvuková apod.). Navržený datový model pro uložení těchto jevů je založen na datovém modelu pro pozorování normy ČSN EN ISO 19156 (Geografická informace – Pozorování a měření), ale byl rozšířen o několik vlastností. Model je samozřejmě dále rozšiřitelný pro jinou geometrickou reprezentaci jevů nebo strukturu atributů. Sbírané jevy jsou ukládány na serveru, odkud jsou dále poskytovány pro další uživatele připojené ke stejné kampani sběru. Všichni uživatelé si mohou sebrané jevy zobrazit nad mapovým podkladem v mobilní aplikaci a případně editovat, pokud mají příslušná oprávnění. Propojení s datovou sadou SPOI je řešeno při pravidelném aktualizacím cyklu hlavního úložiště SPOI. Sebrané jevy v úložišti aplikace SensLog jsou dávkově exportovány v podobě odpovídající datovému modelu SPOI ve formátu RDF/XML.



Obr. 4. HSLCollector – uživatelské rozhraní.

3.4 Vyhledávání dat

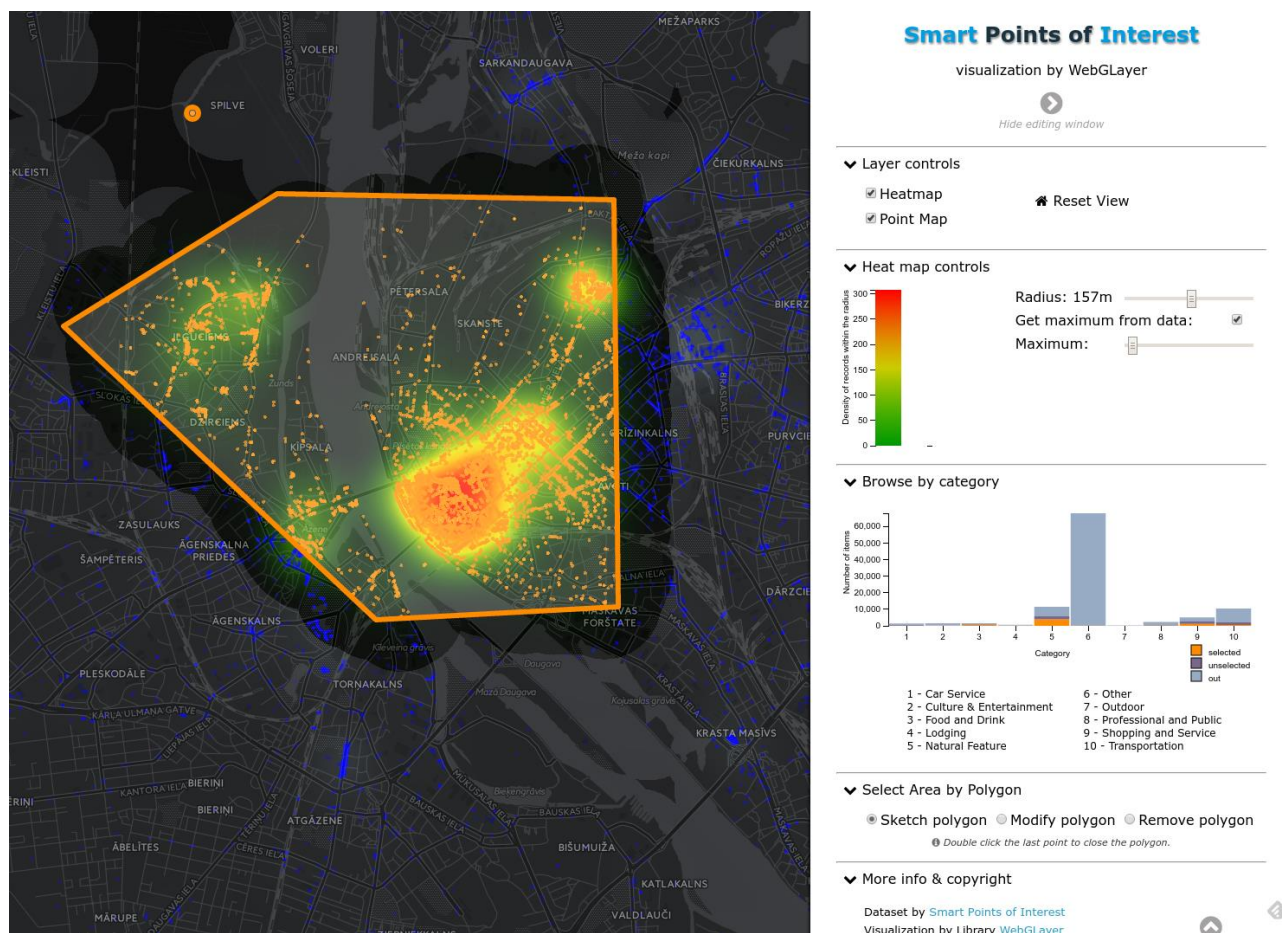
Aplikace nazvaná Smart tourism guide (Obrázek 5) umožňuje vyhledávat body zájmu na určitém území bez znalosti dotazovacího jazyka SPARQL. Uživatel může body získat na základě své současné polohy, zadání názvu sídla (využívána je datová sada GeoNames.org) nebo souřadnic a poloměru prohledávané oblasti. Nalezené body se zobrazí v mapě a uživatel má k dispozici zároveň i jejich seznam, který je možné exportovat do formátů KML a GPX.



Obr. 5. Smart tourism guide.

3.5 WebGLayer vizualizace

Pro hlavní město Lotyšska Rigu byla vytvořena vizualizace SPOI pomocí nástroje WebGLayer (Obrázek 6). Tento přístup je zajímavý především z pohledu možnosti filtrování dat v reálném čase. Realizovaný nástroj zobrazuje data ve formě teplotní mapy (heat map) a umožňuje filtrování na základě polohy (příslušnost ke zvolenému polygonu) a kategorie SPOI. Více informací o WebGLayer řešení je k dispozici například v publikacích Ježek et al. 2015 a Ježek et al. 2017.



Obr. 6. Vizualizace SPOI (WebGLayer).

4 ZÁVĚR

Datová sada SPOI představuje zajímavý a díky svému charakteru do jisté míry i unikátní soubor dat pro tvorbu aplikací zaměřených na cestovních ruch, turismus, dopravu, krizový management a další oblasti. Dá se říci, že pokud chce uživatel svoje data s bodovou geometrií (ale není nutné se omezovat pouze na body, v současnosti vznikají i experimentální datové sady pro linie a plochy) publikovat ve formě propojených dat, představuje SPOI ideální řešení bez ohledu téma dat. V současné době například existují zájemci o využívání SPOI z oborů jako je historie, archeologie, paleontologie nebo precizního zemědělství. V projektu Peregrinus (Západočeská univerzita, Technische Hochschule Deggendorf, Německo a Úhlava o.p.s.) bude na podkladě SPOI vytvářený multimediální průvodce propojující barokní památky a poutní cesty na území jihozápadních Čech a východního Bavorska.

LITERATURA

Berners-Lee, T. (2006). Linked Data - Design Issues. <http://www.w3.org/DesignIssues/LinkedData.html> [cit. 28. 2. 2017].

Bizer, C., Heath, T., & Berners-Lee, T. (2009). Linked data-the story so far. Semantic services, interoperability and web applications: emerging concepts, 205-227.

Čerba, O., Charvát, K., Mildorf, T., Běrzňoš, R., Vlach, P., & Musilová, B. (2016). SDI4Apps Points of Interest Knowledge Base. In Progress in Cartography (pp. 229-237). Springer International Publishing.

Čerba, O., Berzins, R., Charvát, K., & Mildorf, T. (2016). Smart POI: Open and linked spatial data. In EGU General Assembly Conference Abstracts (Vol. 18, p. 12272).

Harris, S., Seaborne, A., & Prud'hommeaux, E. (2013). SPARQL 1.1 query language. W3C recommendation, 21(10).

Heath, T., & Bizer, C. (2011). Linked data: Evolving the web into a global data space. *Synthesis lectures on the semantic web: theory and technology*, 1(1), 1-136.

Ježek, J., Jedlička, K., & Martolos, J. (2015). Visual Analytics of Traffic-Related Open Data and VGI. In *ICIST 2015 Conference*.

Ježek, J., Jedlička, K., Mildorf, T., Kellar, J., & Beran, D. (2017). Design and Evaluation of WebGL-Based Heat Map Visualization for Big Point Data. In *The Rise of Big Spatial Data* (pp. 13-26). Springer International Publishing.