

# Posudek oponenta diplomové práce

Jméno a příjmení autora posudku: Mgr. Petr Škoda

Jméno a příjmení autora práce: Bc. Pavel Hryzlík

Název práce: Využití Linked Data pro sdílení dat o smlouvách veřejných institucí

Cílem práce bylo prozkoumat možnosti využití Linked data v oblasti smluv veřejných institucí. Tento cíl zahrnoval zejména:

- návrh vhodné ontologie pro reprezentaci dat
- konverzi stávajících dat veřejných institucí do podobny Linked data (včetně prolínání)
- návrh a realizace jednotného datového úložiště
- návrh a implementace webové aplikace
- implementaci aplikace, jež by zpřístupnila daná Linked data pro koncové uživatele

Součástí práce bylo napojení na modul Munis ESML, jež je částí informačního systému pro města a obce společnosti Tradia spol. s r.o.

Z hlediska textové podoby je práce vcelku obsáhlá, zahrnuje významné množství diagramů, příkladu použitých dotazů v jazyce SPARQL a tabulek popisujících navržený formát dat a následné mapování na navrženou ontologii. Text diplomové práce je možné rozdělit do dvou částí: část návrhu reprezentace a část popisující platformu pro otevřené smlouvy.

První část začíná nenáročným úvodem do problematiky Linked data. Dále popisuje reprezentaci dat na základě návrhu standardu pro reprezentaci otevřených smluv. Pro tuto reprezentaci pak uvádí neinvazivní přístup konverze (na základě kontextu JSON-LD) do RDF a v rámci možností i do CSV. Samotný popis formátu je podrobný a technicky laděný.

Ve druhé části autor popisuje požadavky na platformu pro otevřená data, její návrh dle těchto požadavků, implementaci a vyhodnocení. Platforma se sestává z uživatelem naprogramované aplikace, ale také nemalého využití existujících nástrojů. Za významnější celky by bylo možné označit:

- Zpřístupnění relačních dat z modulu Munis ESML skrze existující implementaci R2RML.
- Návrh a realizaci pipeline v ETL nástroji UnifiedViews pro sběr dat z katalogizovaných zdrojů.
- Návrh a implementace webové aplikace, jež by zpřístupnila data pro koncového uživatele.

Různorodost použitých programů sice snižuje množství kódu, jež autor musel napsat, na druhou stranu ukazuje schopnost využít různé nástroje pro dosažení definovaných cílů.

Pro úplnost hodnocení je také třeba zmínit, že autor diplomové práce byl součástí skupiny, jež vypracovala použitý návrh standardu a dále se jakožto zaměstnanec firmy Tradia spol. s r.o. podílel na implementaci modulu pro evidenci smluv. Jeho účast na daném tématu tedy překračuje rámec popsaného obsahu samotné diplomové práce.

Pro prezentaci výsledků diplomové práce byla použita uměle připravená data. Využití reálných dat by vyžadovalo souhlas poskytovajícího orgánu veřejné správy a aplikaci anonymizace, jež nebyla součástí diplomové práce – zvolený přístup je tedy přijatelný.

Mezi významnější datové zdroje, jejichž integraci autor prezentoval, patří ARES, RUIAN a CZ DBpedia. Vzhledem k použitým umělým datům, není zcela patrné jakým způsobem a s jakou úspěšností by probíhalo řešení prolínání dat z jednotlivých datových zdrojů. Je však nutné doplnit, že použité techniky pro linkování by se pravděpodobně silně lišily dle očekávané kvality použitých dat. Samotné linkování může být navíc vyřešeno na straně veřejné instituce a pak by nemuselo být součástí platformy.

Celkově je možné konstatovat, že autor splnil zadání diplomové práce.

## Doporučení k obhajobě:

Z výše uvedených důvodů práci doporučuji k obhajobě.

Vynikající práce vhodná pro soutěž studentských prací	ANO <input type="checkbox"/>
-------------------------------------------------------	------------------------------

V Praze dne: 22.1.2015

Podpis: