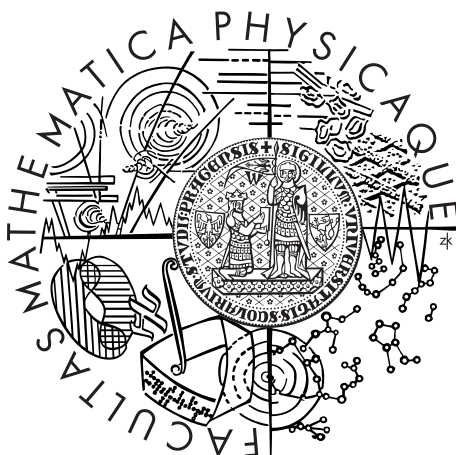


Univerzita Karlova v Praze  
Matematicko-fyzikální fakulta

## DIPLOMOVÁ PRÁCE



Bc. Pavel Hryzlík

# Využití Linked Data pro sdílení dat o smlouvách veřejných institucí

Katedra softwarového inženýrství

Vedoucí diplomové práce: Doc. Mgr. Martin Nečaský, Ph.D.

Studijní program: Informatika

Studijní obor: I2 Softwarové systémy

Praha 2015

Poděkování.

Prohlašuji, že jsem tuto diplomovou práci vypracoval(a) samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů.

Beru na vědomí, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorského zákona v platném znění, zejména skutečnost, že Univerzita Karlova v Praze má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle §60 odst. 1 autorského zákona.

V ..... dne .....

Podpis autora

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

Autor: Bc. Pavel Hryzlík

Katedra: Katedra softwarového inženýrství

Vedoucí diplomové práce: Mgr. Martin Nečaský, Ph.D., Katedra softwarového inženýrství

Abstrakt:

Klíčová slova:

Title: Exploitation of Linked Data for sharing public agreements data

Author: Bc. Pavel Hryzlík

Department: Department of Software Engineering

Supervisor: Doc. Mgr. Martin Nečaský, Ph.D., Department of Software Engineering

Abstract:

Keywords:

# Obsah

<b>1</b>	<b>Úvod</b>	<b>3</b>
1.1	Motivace . . . . .	4
1.2	Cíl práce . . . . .	5
1.3	Struktura práce . . . . .	5
<b>2</b>	<b>Otevřená data a principy Linked Data</b>	<b>6</b>
2.1	Otevřená data (Open Data) . . . . .	6
2.2	Kvalita otevřených dat . . . . .	8
2.3	Stupně otevřenosti <sup>1</sup> . . . . .	8
2.4	Propojitelná data (Linked Data) . . . . .	10
2.5	Otevřená a propojitelná data (Linked Open Data - LOD) . . . . .	11
2.6	Výhody a přínosy otevřených dat a principů Linked data . . . . .	12
2.7	RDF (Resource Description Framework) . . . . .	13
2.8	RDFOntologie . . . . .	14
2.8.1	Propojování se souvisejícími entitami . . . . .	16
2.9	Publikace . . . . .	17
2.9.1	Příklad dat serializovaných ve formátu N-Triples . . . . .	17
2.9.2	Příklad dat serializovaných ve formátu Turtle . . . . .	18
2.9.3	Příklad dat serializovaných ve formátu JSON-LD . . . . .	19
<b>3</b>	<b>Otevřené smlouvy</b>	<b>22</b>
3.1	Situace ve veřejné správě ČR . . . . .	22
3.2	Standard pro zveřejňování smluv . . . . .	23
3.2.1	Základní struktura . . . . .	23
3.2.2	Reprezentované entity . . . . .	26
3.2.3	Číselníky . . . . .	32
3.3	Publikace . . . . .	33
3.3.1	JSON . . . . .	33
3.3.2	CSV . . . . .	37
3.4	Metodika zveřejňování smluv . . . . .	38
<b>4</b>	<b>Otevřené smlouvy jako Linked Data</b>	<b>39</b>
4.1	Přiřazení identifikátorů jednotlivým entitám otevřených smluv . . . . .	39
4.2	Ontologie pro publikaci dat o smlouvách . . . . .	40
4.2.1	Analýza vhodných, již existujících ontologií . . . . .	41
4.2.2	Tvorba ontologie . . . . .	43
4.2.3	Publikace . . . . .	48
4.3	Možnosti propojení na související data . . . . .	48
4.4	Provázání s datovým formátem JSON . . . . .	49
<b>5</b>	<b>Požadavky na platformu pro otevřené smlouvy</b>	<b>55</b>
5.1	Funkční požadavky . . . . .	55
5.2	Nefunkční požadavky . . . . .	55

<b>6</b>	<b>Návrh platformy pro otevřené smlouvy</b>	<b>56</b>
6.1	Architektura . . . . .	56
6.1.1	Konverzní mechanismus . . . . .	56
6.1.2	Jednotné úložiště . . . . .	56
6.1.3	Webová aplikace . . . . .	56
6.2	Linked Data v procesu otevírání smluv . . . . .	56
<b>7</b>	<b>Implementace platformy</b>	<b>57</b>
7.1	Konverzní mechanismus . . . . .	57
7.1.1	Munis ESML . . . . .	57
7.1.2	R2RML mapování . . . . .	57
7.1.3	Volba R2RML procesoru . . . . .	57
7.1.4	Volba technologií a implementační platformy . . . . .	57
7.1.5	Napojení na datové úložiště . . . . .	57
7.1.6	SPARQL endpoint . . . . .	57
7.1.7	Zpracování RDF výstupu . . . . .	57
7.2	Jednotné úložiště . . . . .	57
7.2.1	Nástroj Unified views . . . . .	57
7.3	Webová aplikace . . . . .	57
<b>8</b>	<b>Zátěžové testy</b>	<b>58</b>
8.1	Předpokládaná situace na úřadech . . . . .	58
8.2	Výsledky . . . . .	58
	<b>Závěr</b>	<b>59</b>
	<b>Seznam použité literatury</b>	<b>60</b>
	<b>Seznam tabulek</b>	<b>61</b>
	<b>Seznam použitých zkratk</b>	<b>62</b>
	<b>Přílohy</b>	<b>63</b>
	<b>A Příloha</b>	<b>64</b>
	<b>B Příloha</b>	<b>65</b>
	<b>C Příloha</b>	<b>66</b>
	<b>D Příloha</b>	<b>67</b>

# 1. Úvod

V době informační společnosti se využívání internetu stalo naší každodenní rutinou. Skrze různé webové aplikace a služby každodenně pracujeme s obrovským množstvím informací. Běžně komunikujeme přes e-mail, finance spravujeme skrze internetové bankovníctví, část svého osobního života sdílíme na sociálních sítích. Požadavek na on-line vyřizování agendy vůči veřejné správě tedy není překvapujícím.

Problematika elektronizace veřejné správy, jednotně nazývaná jako „e-government“, je aktuálním tématem již po mnoho let. Důsledkem tohoto procesu je generování obrovského množství nesmírně důležitých dat. Tato data ale v naprosté většině případů leží schovaná v databázích jednotlivých veřejných institucí. Mnoho z těchto dat by ale ze zákona mělo být volně dostupných. Často však jediným možným způsobem, jak taková data získat je použití zákona č.106/1999 Sb.<sup>1</sup>, o svobodném přístupu k informacím. Netřeba zmiňovat, že tato snaha se mnohdy může stát značně netriviální.

Řešením je vhodná data, resp. metadata o těchto datech, zpřístupnit on-line. Pro strojově čitelná data zveřejněná na internetu se zažil pojem Otevřená data. Tato data pak může vyhledávat a zpracovávat kdokoli. To přináší řadu dílčích výhod od úspory nákladů, přes boj s korupcí, až po zapojení občanů, nemluvě o podnikatelském potenciálu, převážně možnosti vzniku mnoha užitečných aplikací pracujících nad otevřenými daty. To celé za cenu minimálních nákladů z veřejných rozpočtů.

Otvírání dat můžeme chápat jako další krok v procesu elektronizace veřejné správy. Průkopníky v této oblasti jsou státy s vyspělou formou demokracie, jako USA a Spojené království. Příklad si ale také můžeme vzít od Estonska. Malá země, vědoma si, že nemá nerostné bohatství ani rozvinutý průmysl, se rozhodla prosadit na poli informačních technologií, kde základem jsou otevřené on-line služby veřejné správy. Důležitost otevřených dat si uvědomuje i Evropská unie. Směrnicí 2013/37/EU<sup>2</sup> v podstatě doporučuje členským státům, aby data otvíraly. České republice se také povedlo nastartovat procesy otvírání veřejné správy. Pokrok je cítit hlavně na národní úrovni. Mezi městy a obcemi jsou však otevřená data často stále neznámým pojmem. Problematikou a obecně osvětou otevřených dat se zabývá mimo jiné Ministerstvo vnitra ČR<sup>3</sup>, projekt Rekonstrukce státu<sup>4</sup>, Fond Otakara Motejla<sup>5</sup>, Oživení o.s.<sup>6</sup> či iniciativa OpenData.cz<sup>7</sup>.

Otevřená data však nelze chápat jako samospásné řešení problémů veřejné správy. Jsou spíše prostředkem ke zvýšení otevřenosti a transparentnosti. Veřejná služba však může být netransparentní i s otevřenými daty. Řekněme, že pro kvalitní veřejnou službu jsou otevřená data nutnou, nikoli však postačující podmínkou.

Dalším aspektem otevřených dat je jejich kvalita. Kvalitní otevřená data jsou

---

<sup>1</sup>Zákon č.106/1999 Sb. - <http://www.zakonyprolidi.cz/cs/1999-106>

<sup>2</sup>Směrnice 2013/37/EU - <http://www.eurlex.cz/dokument.aspx?celex=32013L0037>

<sup>3</sup>Ministerstvo vnitra - <http://www.mvcr.cz/>

<sup>4</sup>Rekonstrukce státu - <http://www.rekonstrukcestatu.cz/cs/>

<sup>5</sup>Fond Otakara Motejla - <http://www.motejl.cz/>

<sup>6</sup>Oživení o.s. - <http://www.oziveni.cz/>

<sup>7</sup>OpenData.cz - <http://opendata.cz/>

propojena mezi sebou v rámci jednotného sdíleného prostoru, mohou na sebe odkazovat a využívat širokého kontextu, které takový sdílený prostor propojených dat nabízí. Taková data využívají principů Linked Data.

## 1.1 Motivace

Základní motivaci pro vznik této práce bych rozdělil do tří pilířů:

### Veřejnoprávní sféra

Na podzim roku 2014 se konal seminář Transparentnost v obcích v Poslanecké sněmovně pořádaný panem Mgr. Janem Farským. V rámci semináře se sešla skupina složená ze zástupců měst a obcí, akademické sféry a neziskového sektoru. Předmětem jednání byla otevřená data. Výsledkem bylo rozhodnutí, že první datovou sadou vhodnou k plošnému otevření, také vzhledem k chystanému zákonu o registru smluv, jsou údaje o smlouvách. Prvním krokem je standardizace datového formátu, resp. určení položek vhodných ke zveřejnění. Motivací bylo, že pokud standard začne využívat netriviální počet měst a obcí, tak je reálná šance k prosazení standardu na národní úroveň. Ustanovila se tedy, pod záštitou Oživení o.s. a Centra aplikované ekonomie o.s.<sup>8</sup>, „akční“ skupina, jejímž cílem byla tvorba datového standardu pro otevřené smlouvy. Bylo mi ctí, se stát členem této skupiny.

### Komerční sféra

Jako externista se podílím na tvorbě software pro veřejnou správu ve společnosti Triada spol. s r.o. Mým úkolem se ke konci roku 2014 stala tvorba modulu ESML pro interní evidování smluv.

### Akademická sféra

V rámci MFF UK ve spolupráci s Fakultou informatiky VŠE vznikla iniciativa OpenData.cz. Jejím cílem je vybudování otevřené datové infrastruktury v České republice. Na MFF UK také probíhá výzkum propojitelných dat, Linked Data. Mým cílem bylo přispět k otevřené datové infrastruktuře, navíc s využitím principů Linked Data. Rozhodnutí věnovat se publikaci dat o smlouvách padlo již v červnu 2014. Konkrétní obrysy však práce získala až s přispěním výše zmíněných pilířů. Výsledkem je tedy aplikace principů Linked Data pro publikaci a sdílení dat o smlouvách s možností konkrétního využití nad modulem ESML společnosti Triada. To celé s ohledem na vznikající datový standard. Jednou z důležitých motivací bylo, že v případě prosazení datového standardu na národní úroveň mohou města a obce používající modul ESML využitím této práce automaticky zveřejňovat smlouvy v Linked Data podobě, a to s minimálními náklady. Taková data lze pak agregovat do jednotných úložišť, nad kterými mohou vznikat nejrůznější aplikace přinášející konečný přínos pro uživatele.

---

<sup>8</sup>Centrum aplikované ekonomie o.s. - <http://cae.zindex.cz/>



## 1.2 Cíl práce

Cílem práce je prozkoumat možnosti využití principů Linked Data pro publikaci a sdílení dat o smlouvách veřejných institucí a jejich propojení na související data ve veřejném prostoru. Prvním krokem je definování datového standardu a ontologie pro otevřené smlouvy. Dalším krokem je návrh způsobu konverze dat stávajícími informačními systémy veřejných institucí (v podobě relačních databází) do otevřeného formátu využívající principy Linked Data a implementace konverzního mechanismu pro vybraný konkrétní informační systém (Triada spol. s.r.o.). V dalším kroku následuje návrh a implementace jednotného úložiště dat o smlouvách v Linked Data s experimentálním zprovozněním na serveru poskytnutém vedoucím práce. V jednotném úložišti se očekává návrh řešení integračních problémů dané heterogenitou dat publikovaných různými veřejnými institucemi. Následujícím krokem je nad tímto jednotným úložištěm návrh a implementace webové aplikace, která data o smlouvách zpřístupní koncovým uživatelům.

## 1.3 Struktura práce

Obsah práce je rozdělen na 9 kapitol a 3 přílohy. Ve druhé kapitole jsou popsány a vysvětleny základní principy otevřených dat. Třetí kapitola se zabývá pojmem otevřené smlouvy. Kapitola nejdříve rozebere aktuální stav otevřenosti smluv ve veřejné správě a následně nastíní vznikající datový standard. Čtvrtá kapitola zadefinuje otevřené smlouvy jako Linked data. V páté kapitole se analyzují požadavky na platformu pro otevřené smlouvy. Šestá kapitola zmíněnou platformu navrhne. Sedmá kapitola se zabývá konkrétní implementací platformy. V osmé kapitole jsou znázorněny zátěžové testy některých dílčích částí implementace. Poslední, devátou kapitolou je závěr shrnující práci jako celek. Nedílnou součástí práce je seznam použité literatury a slovníček pojmů. Práce zahrnuje také 3 přílohy. V příloze A je znázorněn harmonogram vývoje standardu otevřených smluv. V příloze B se nachází uživatelská dokumentace. Konečně, příloha C popisuje strukturu přiloženého datového nosiče.

## 2. Otevřená data a principy Linked Data

Předmětem této kapitoly je čtenáře stručně seznámit se základními pojmy a principy otevřených, propojitelných dat a následně s technologiemi sloužícími k jejich zápisu a zpracování.

### 2.1 Otevřená data (Open Data)

*„Open data can help us address the greatest challenges of our time and generate value for everyone“* - Open Data Institute 2012

Začneme definicí, kterou si postupně vysvětlíme. Jako otevřená data můžeme chápat údaje zveřejněná na internetu, která jsou[ZDROJ]:

1. úplná
2. snadno dostupná
3. strojově čitelná
4. používající standardy s volně dostupnou specifikací
5. zpřístupněna za jasně definovaných podmínek užití dat s minimem omezení
6. dostupná uživatelům při vynaložení minima možných nákladů

#### Úplnost

Pokud se rozhodneme zveřejňovat data, tak v případě že nás neomezuje zákon, či jiná restriktivní opatření, měli bychom dbát na to, aby byla úplná, resp. v maximálním možné rozsahu. Není cílem zveřejňovat útržky ztrácející vypovídající hodnotu.

#### Snadná dostupnost

Základní požadavek na dostupnost otevřených dat spočívá v tom, že by měla být k dispozici kdykoli, ne pouze např. na vyžádání. Otevřená data budou také přínosem pro širokou veřejnost jedině tehdy, pokud budou snadno dohledatelná. Skrytá data za změtí odkazů se hledají špatně.

#### Strojová čitelnost

Klíčovou vlastností otevřených dat je strojová čitelnost. Otevřeným datům by měl porozumět nejen člověk, ale i stroj. Účelem je umožnit data automatizovaně zpracovávat, analyzovat, počítat statistiky apod.

## Otevřené standardy

Software, nástroje či metodiky potřebné k zpracování dat by měly být volně dostupné. Data v uzavřeném formátu, které potřebují ke zpracování konkrétní proprietární software, postrádají smysl otevřenosti.

## Zpřístupněna za jasně definovaných podmínek

Typicky je třeba dbát nato, aby data byla zveřejňována pod otevřenou licenci.<sup>1</sup>

## Dostupná uživatelům s minimem nákladů

Je třeba si uvědomit, že nezveřejňujeme data pro data. Zveřejňujeme pro přidanou hodnotu, např. pro lepší službu nebo vyšší efektivitu. Náklady na zveřejnění by tak neměly přesáhnout případná zlepšení.



Obrázek 2.1: Logo otevřených dat

---

<sup>1</sup>Více k problematice licencování a užití otevřených dat lze dohledat na webu Ministerstva vnitra - <http://www.mvcr.cz/clanek/otevrena-data.aspx?q=Y2hudW09Mg%3d%3d>

## 2.2 Kvalita otevřených dat

Tvůrce WWW a ředitel konsorcia W3C<sup>2</sup> Tim Berners-Lee<sup>3</sup> navrhl pěti hvězdičkový systém, jak kategorizovat otevřená data (viz Obrázek 2.2). Každá hvězdička definuje stupeň otevřenosti, kde 5★ znamená nejvyšší kvalitu dat, 1★ naopak nejmenší. Také platí, že každý stupeň je nadmnožinou (rozšíření) stupně předešlého.

## 2.3 Stupně otevřenosti<sup>4</sup>

### ★ Libovolná zveřejněná data pod otevřenou licencí

- Přínosy pro uživatele - uživatel může data číst, tisknout, ukládat, přenášet, měnit a sdílet podle svého uvážení
- Přínosy/náklady pro vydavatele - velmi nenáročné na publikaci
- Příkladem může být formát PDF

Publikace dat na úrovni 1★ je zdaleka nejjednodušší a nepotřebuje příliš vynaloženého úsilí. Určitě je lepší zveřejňovat data na úrovni 1★, než vůbec. Využitelnost dat však může být velmi obtížná, např. díky nutnosti dolování dat z PDF dokumentů.

### ★★ Strukturovaná data ve strojově čitelném formátu

- Přínosy pro uživatele - uživatel může pokročile zpracovávat data s využitím proprietárních nástrojů k tomu určených
- Přínosy/náklady pro vydavatele - velmi nenáročné na publikaci
- Příkladem může být formát MS Excel (.xls)

V dnešní době už poměrně rozšířený způsob publikace dat. Zpracování dat ale vyžaduje specifické nástroje k tomu určené. Pokud tedy chceme zpracovávat např. excelovskou tabulku (.xls), potřebujeme k tomu komerční produkt MS Excel<sup>5</sup>.

### ★★★ Formát dat je otevřený

- Přínosy pro uživatele - uživatel při zpracování dat není omezen žádným specifickým nástrojem
- Přínosy/náklady pro vydavatele - nenáročné na publikaci, může však vyžadovat transformaci dat, např. z uzavřeného formátu

---

<sup>2</sup>W3C - <http://www.w3.org/>

<sup>3</sup>Tim Berners-Lee - životopis - <http://www.w3.org/People/Berners-Lee/>

<sup>4</sup>5 star ifno - <http://5stardata.info/>

<sup>5</sup>Toto se netýká formátu .xlsx. Ten již vychází z otevřené specifikace Office Open XML - ata publikovaná v .xlsx formátu tedy můžeme chápat jako 3★.

- Příkladem může být formát CSV

Teprve v této kategorii se můžeme bavit o “opravdových” otevřených datech. Resp. data musejí mít stupeň otevřenosti minimálně 3★ , aby naplnila základní definici otevřených dat uvedenou výše.

## ★★★★ Jednotlivé objekty jsou identifikovány pomocí URI

- Přínosy pro uživatele - uživatel se může na data odkazovat, odkazy si ukládat, případně data snadno kombinovat s jinými (na stejném, nebo vyšším stupni)
- Přínosy/náklady pro vydavatele - náročnější na publikaci
- Příkladem může být formát RDF

Důležité je dbát na to, aby URI nebylo virtuální, resp. po dotázání se uživateli vrátil požadovaný obsah. V prostředí WWW je zajištění obsahu typicky praktikováno skrze protokol HTTP.

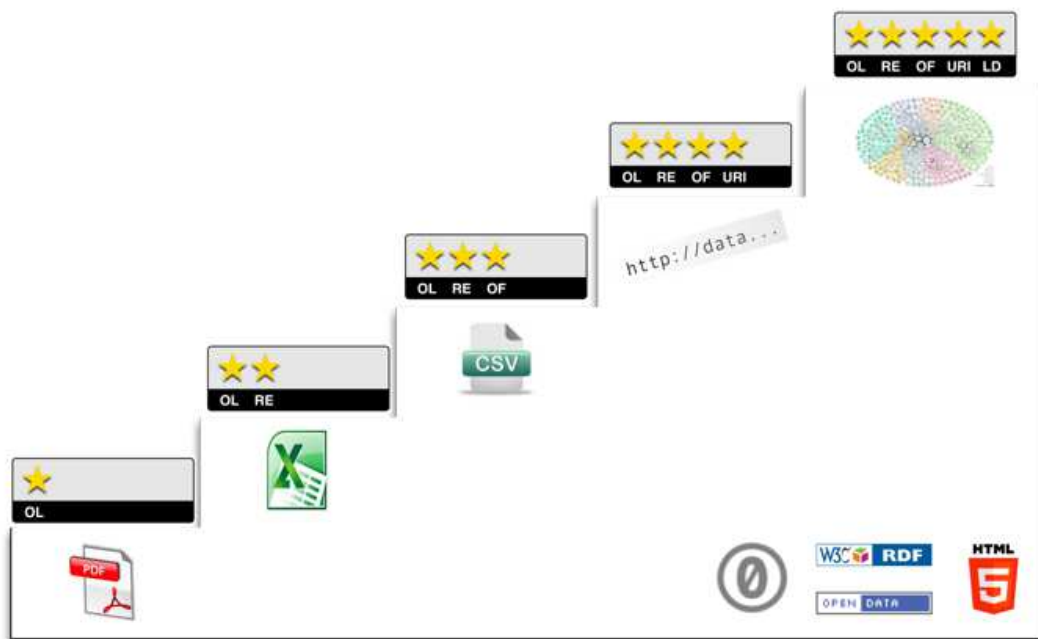
Díky URI identifikaci můžeme data reprezentovat jako orientovaný graf propojených objektů, které na sebe mohou vzájemně odkazovat. K popisu takovýchto dat se používá formát RDF.[ODKAZ NA KAPITOLU]

V prostředí České republiky považováno jako nadstandard.

## ★★★★★ Data jsou propojena se souvisejícími daty

1. Přínosy pro uživatele - vytvoření efektu datové sítě, větší informační hodnota dat
2. Přínosy/náklady pro vydavatele - náročnější na publikaci
3. Příkladem může být formát RDF

V této nejvyšší kategorii se data mohou stát součástí datové sítě propojených grafů.



Obrázek 2.2: Stupně otevřenosti dat, zdroj: <http://5stardata.info/>

## 2.4 Propojitelná data (Linked Data)

Linked Data vychází z myšlenky webu aplikované na data. Webu rozumíme jako síti propojených webových stránek. Cílem Linked Data je mít síť propojených, strojově čitelných dat, resp. stavební kámen sémantického webu<sup>6</sup>. Jedná se v podstatě o další krok v evolučním vývoji webu jako takového.

Podle [7] definujeme základní principy Linked Data jako:

1. Každá entita je identifikována pomocí HTTP URI<sup>8</sup>
2. HTTP URI by mělo být vyhledatelné v síti WWW a umožňovat k němu přistupovat a odkazovat se na něj
3. Po přistoupení na HTTP URI entity mají být poskytnuty relevantní informace o dané entitě ve standardizovaném formátu či prostřednictvím API<sup>9</sup>
4. Data k entitám rozšířit o HTTP URI odkazy na další související entity<sup>10</sup>

Jak je vidět, Linked Data naplňují všechny požadavky na 5★ kvalitu dat s jednou výjimkou. Linked Data nemusejí být z podstaty otevřenými daty. Určitě si dovedeme představit mnoho scénářů, kdy je přínosem mít propojená, ale privátní data. Typickým příkladem můžou být korporátní intranetové informační systémy.

<sup>6</sup>Semantic web - <http://www.cs.umd.edu/~golbeck/LBSC690/SemanticWeb.html>

<sup>7</sup>Linked Data - <http://www.w3.org/DesignIssues/LinkedData.html>

<sup>8</sup>V prostředí webu to bude typicky hypertextový odkaz ve formě URL[ZDROJ]

<sup>9</sup>Pro popis Linked Data se typicky používá jazyk RDF - <http://www.w3.org/RDF/>, k dotazování k datovému API - SPARQL - <http://www.w3.org/TR/rdf-sparql-query/>

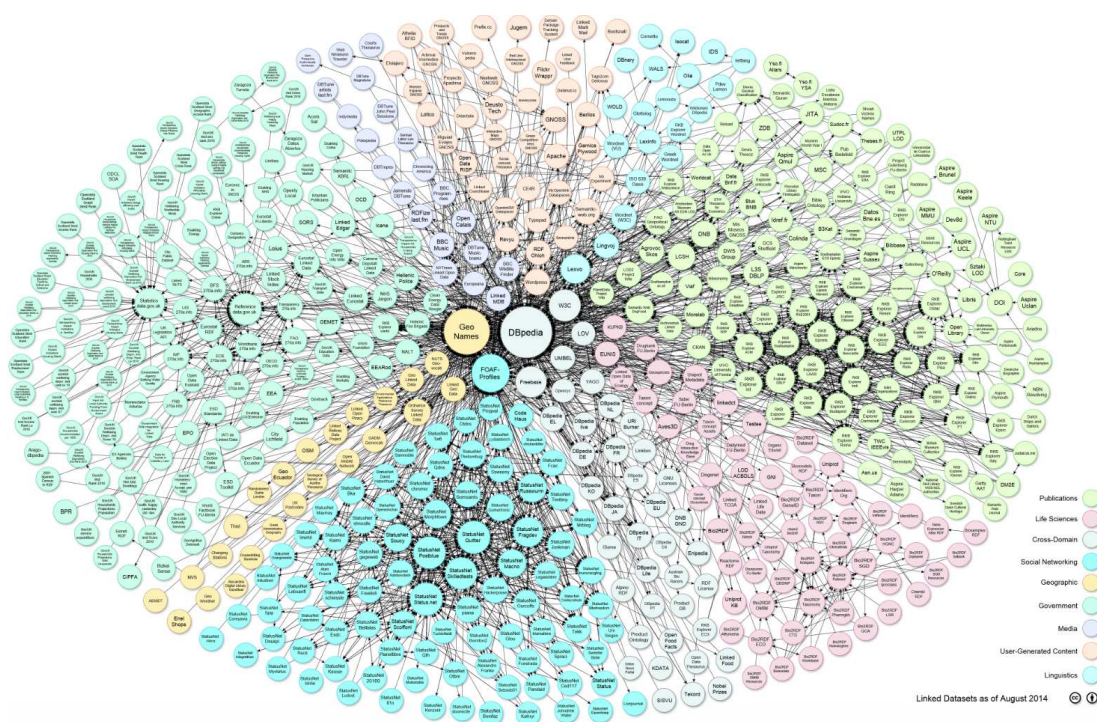
<sup>10</sup>o nám zaručí, že můžeme procházet jednotlivé entity podobným způsobem jako webové stránky v rámci sítě WWW.

## 2.5 Otevřená a propojitelná data (Linked Open Data - LOD)

Otevřená data na úrovni 5★ kvality můžeme tedy chápat jako Linked Open Data. Taková data se mohou stát součástí globálního prostoru sdílených, propojených dat. Připojením datové sady tak můžeme čerpat informační potenciál celého prostoru<sup>11</sup>.

Takový prostor s časem neustále roste. Využití lze nalézt ve většině oblastí lidského konání. Od sdílení a obohacování vědeckých dat, např. biologických, chemických struktur a reakcí s cílem objevů nových postupů v medicíně, přes zpracování dat jednotlivých veřejných správ za účelem kvalitnější veřejné služby až po obohacování kontextu nejrůznějšího mediálního obsahu.

Na obr. 2.3 vidíme příklad vizualizace otevřených a propojených (LOD) dat nazývaný Linked Open Data Cloud. Jedná se o datasety obsahující alespoň 1000 trojic (více v kapitole o RDF) a alespoň 50 odkazů na jiná data ve sdíleném prostoru.



Obrázek 2.3: Linked Open Data Cloud, Srpen 2014, zdroj: <http://lod-cloud.net/>

<sup>11</sup>Tvůrci grafu procházejí web a do cloudu přidávají dostupné datasety splňující podmínky Linked Data a podmínky na počet trojic a odkazů. Nezkontrolují ale licence jednotlivých datasetů. Některé datasety proto mohou být chráněny specifickými právy.

## **2.6 Výhody a přínosy otevřených dat a principů Linked data**

### **Obecné výhody otevřených dat**

1. Zapojení uživatelů - kontrola, návrhy ke zlepšení dat
2. Zvýšení transparentnosti vydavatele dat, boj s korupcí
3. Kvalitnější veřejná služba
4. Zvýšení efektivity, úspora nákladů
5. Široké možnosti dalšího využití - analýzy, statistiky, vizualizace

### **Výhody principů Linked Data**

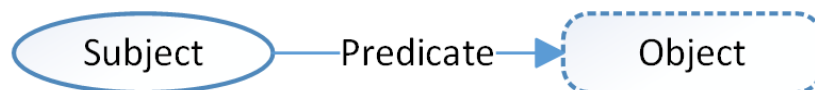
1. Sdílená, rozšiřitelná a snadno znovu použitelná data
2. Data jsou začleněna do kontextu, resp. lze se odkazovat přímo na data
3. Data jsou propojena s dalšími relevantními daty, informační hodnota dat je tedy tím Větší, čím více mají vazeb
4. Standardizované formáty pro publikaci



## 2.7 RDF (Resource Description Framework)

Formát RDF byl vyvinut za účelem snadného strojového zpracování a propojování dat. Jedná se o čistě abstraktní formát udávající, jak data popisovat. Nezabývá se tedy konkrétní podobou výsledných dat.

Základním stavebním kamenem RDF je tvrzení, resp. trojice: **Subjekt - Predikát - Objekt** (viz Obrázek 2.4). Subjektem je míněn zdroj, který popisujeme. Predikát je vlastnost, která o objektu něco tvrdí. Objekt je hodnota dané vlastnosti. Jednotlivé trojice mohou na sebe navazovat a vytvořit tak orientovaný graf.



Obrázek 2.4: Základní RDF trojice

### Nyní definujeme několik pravidel a doporučení pro popisování dat v RDF

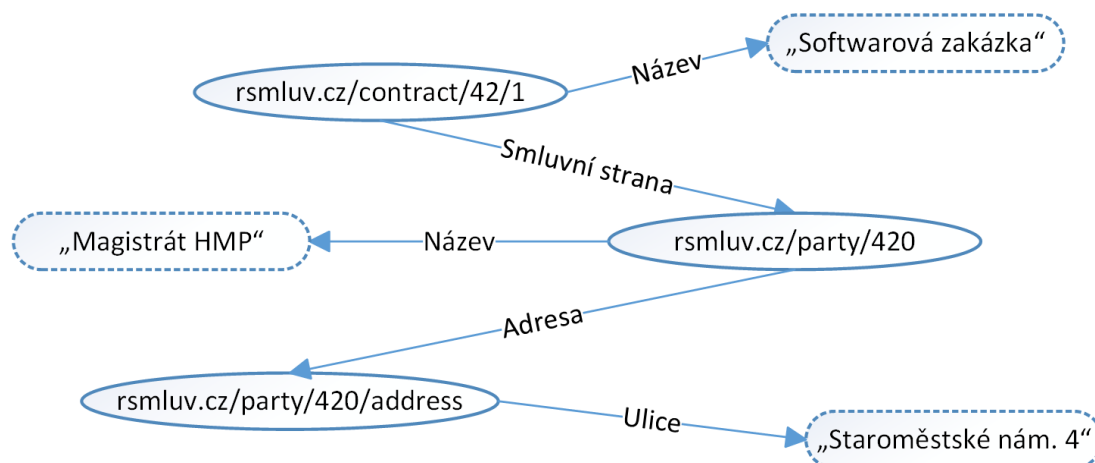
1. Každý subjekt je jednoznačně identifikován pomocí URI, nebo je označen jako anonymní<sup>12</sup>
2. Objektem je buď hodnota (literál), odkaz na subjekt (resource), nebo je označen jako anonymní
3. Pro každý subjekt je specifikován jeho typ (třída) formou URI
4. Každému predikátu je přiřazen také jeho typ formou URI
5. Jednotlivé URI z bodů 4,5 by měly odkazovat na konkrétní slovníky tříd a predikátů, resp. ontologie

Na obr. 2.5 vidíme příklad jednoduchého grafu ve formátu RDF (aplikována pravidla 1 a 2). Popisuje 3 subjekty a přiřazuje jim konkrétní vlastnosti. Vidíme, že každý subjekt je identifikován vlastním URI. Díky tomu mohou subjekty na sebe odkazovat. Jednotlivé trojice by pak vypadaly takto:

1. *http://rsmluv.cz/contract/42/1 - Název - Softwarová zakázka*
2. *http://rsmluv.cz/contract/42/1 - Smluví strana - rsmluv.cz/party/420*
3. *http://rsmluv.cz/party/420 - Název - Magistrát HMP*
4. *http://rsmluv.cz/party/420 - Adresa - rsmluv.cz/party/420/address*
5. *http://rsmluv.cz/party/420/address - Ulice - Staroměstské nám. 4*

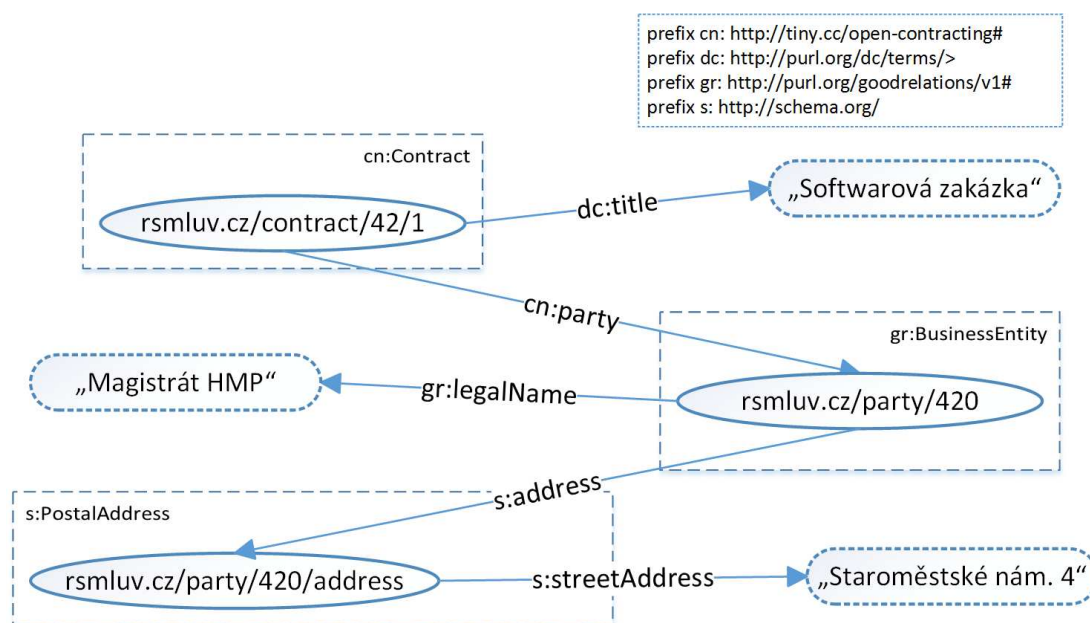
---

<sup>12</sup>Subjekty, příp. objekty lze označit jako anonymní, resp. Blank node. Na anonymní subjekty, resp. objekty ale nelze přímo přistupovat. Používají se typicky k zapouzdření, či jako kontejnery jiných objektů.



Obrázek 2.5: Jednoduchý RDF graf

Ze zmíněného příkladu ale není zřejmý význam, resp. sémantika jednotlivých subjektů a predikátů. Je tedy důležité jim přiřadit konkrétní typy. Každý typ by měl být popsán v konkrétním slovníku tříd a predikátů. Takovéto slovníky nazýváme ontologiemi. Na obr. 2.6 vidíme zmíněný příklad rozšířený o přiřazené typy (aplikována pravidla 3, 4, 5)<sup>13</sup>.



Obrázek 2.6: RDF graf s přiřazenými typy

## 2.8 RDFOntologie

Pod pojmem ontologie si můžeme představit sadu termínů popisujících určitou věcnou oblast. V případě popisování RDF dat definujeme slovník tříd a vlastností (predikátů), které mohou uživatelé ve svých datech používat.

<sup>13</sup>Pro zapisování typů se kvůli úspornosti používají prefixy definované typicky na začátku dokumentu.

Konkrétní ontologii nelze chápat jako striktně vyžadovaný standard, ale spíše jako sadu doporučení. Buď využijeme k popisu dat nějakou z řady již existujících ontologií, nebo můžeme vytvořit ontologii vlastní. Přesto ale chceme, aby se již existující ontologie používaly co nejvíce. Přínosem je hlavně to, že aplikace a nástroje implementované nad známými ontologiemi budou schopné automaticky rozpoznat naše data<sup>14</sup>.

Základními jazyky pro modelování RDF dat jsou Web Ontology Language (OWL)<sup>15</sup> a RDF Schema (RDFS)<sup>16</sup>. Konkrétní specifikace se provádí opět ve formátu RDF a je publikována pod vlastním URI.

Mezi základní výrazové prostředky jazyka OWL a RDFS patří:

- *owl:Class* - typ entity třída
- *owl:ObjectProperty* - typ entity vlastnost
- *owl:FunctionalProperty* - typ jedinečná vlastnost (nemůže se opakovat)
- *owl:unionOf* - jeden typ třídy z výčtu musí být vyplněn
- *owl:equivalentClass* - definuje, že se jedná o třídu odpovídající jiné třídě
- *owl:equivalentProperty* - definuje, že se jedná o vlastnost odpovídající jiné vlastnosti
- *rdfs:label* - popis třídy/vlastnosti
- *rdfs:comment* - komentář třídy/vlastnosti
- *rdfs:domain* - požadovaný typ domény třídy/vlastnosti
- *rdfs:range* - požadovaný rozsah typů třídy/vlastnosti
- *rdfs:isDefinedBy* - definice zdroje třídy/vlastnosti
- *rdfs:subClassOf* - definice, že se jedná o podtřídu určité třídy
- *rdfs:subPropertyOf* - definice, že se jedná o podvlastnost určité vlastnosti

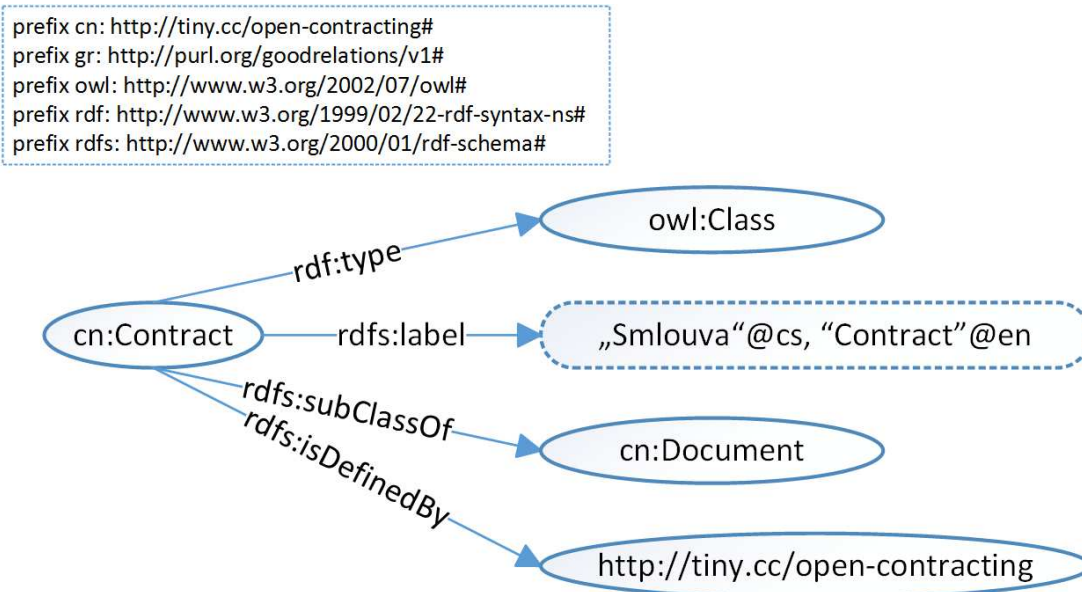
Na obr. 2.7 vidíme příklad ontologie třídy Contract. Ontologie nám říká, že se jedná o třídu (typ *owl:Class*) s názvem Smlouva (*rdfs:label*), která je podtřídou (*rdfs:subClassOf*) Document a je definovaná (*rdfs:isDefinedBy*) v ontologii <http://tiny.cc/open-contracting>. Kdokoli pak bude zpracovávat entitu označenou tímto typem, tak díky přiřazené ontologii bude schopen určit, že se jedná o smlouvu.

---

<sup>14</sup>Mezi všeobecně známé ontologie patří např. DublinCore - <http://purl.org/dc/terms/>, Friend-of-a-Friend - <http://xmlns.com/foaf/0.1/> nebo Schema - <http://schema.org/>. Existuje také katalog ontologií - <http://lov.okfn.org/dataset/lov/>

<sup>15</sup>Web Ontology Language - <http://www.w3.org/2001/sw/wiki/OWL>

<sup>16</sup>RDF Schema - <http://www.w3.org/TR/rdf-schema/>



Obrázek 2.7: Ontologie třídy *Contract*

### 2.8.1 Propojování se souvisejícími entitami

Díky RDF můžeme data reprezentovat jako orientovaný graf. Otázka tedy zní, zda-li lze propojovat grafy mezi sebou. Ve formátu RDF je to velmi jednoduché. Jako objekt predikátu stačí položit subjekt z jiného grafu. Díky URI identifikaci entit tedy není rozdílem, zda-li je cílovým subjektem entita v lokálních datech, nebo entita cizí.

V rámci propojování dat s jinými datasety však není neobvyklé, že stejné entity jsou reprezentované v různých datasetech pod vlastními URI. Je tedy třeba vyjádřit, že se jedná o data reprezentující stejné entity. V jazyku OWL za tímto účelem existuje predikát `owl:sameAs`, kterým můžeme definovat odpovídající si entity (viz Obrázek 2.8).



Obrázek 2.8: Možnost propojení dat

## 2.9 Publikace

V minulých kapitolách bylo řečeno, jak popisovat data pomocí RDF. Jednalo se o sémantický popis. Pokud však data chceme publikovat, je třeba konkrétního datového formátu, který definuje syntaxi, resp. jak RDF data serializovat. Takových formátů existuje celá řada, např.

- **N-Triples**<sup>17</sup> - nejjednodušší serializace RDF grafu v podobě výčtu trojic
- **N-Quads**<sup>18</sup> - rozšíření pro N-Triples s možností zaznamenat více grafů
- **RDF/XML**<sup>19</sup> - RDF graf serializovaný do XML, využívající prefixového zápisů
- **Turtle**<sup>20</sup> - úsporný textový formát s možností komprese trojic, využívající prefixových zápisů
- **Trig**<sup>21</sup> - rozšíření Turtle pro použití nad více grafy
- **RDFa**<sup>22</sup> - serializace RDF do (X)HTML<sup>23</sup> dokumentů, využívající prefixového zápisů
- **JSON-LD**<sup>24</sup> - specifický zápis RDF grafu, využívající mapování položek JSON dokumentu na RDF ontologie

Pro potřeby této práce si vystačíme s formáty N-Triples, Turtle a JSON-LD. Vysvětlíme si je na příkladech. Jako data k serializaci použijeme příklad z obr. 2.6.

### 2.9.1 Příklad dat serializovaných ve formátu N-Triples

Serializace RDF dat do N-Triples je velmi jednoduchá. Jedná se o seznam trojic oddělených tečkou. Každá trojice je uvedena na vlastním řádku. Tento formát nepoužívá prefixové zkracování URI. Je vhodný pro proudové zpracování velkého množství dat (viz Obrázek 2.1)<sup>25</sup>.

```
1 <http://rsmluv.cz/contract/42/1>
2   <http://www.w3.org/1999/02/22-rdf-syntax-ns#type>
3     <http://tiny.cc/open-contracting#Contract> .
4 <http://rsmluv.cz/contract/42/1>
5   <http://purl.org/dc/terms/title>
6     "Softwarová zakázka" .
7 <http://rsmluv.cz/contract/42/1>
8   <http://tiny.cc/open-contracting#party>
```

<sup>17</sup>N-Triples - <http://www.w3.org/TR/n-triples/>

<sup>18</sup>N-Quads - <http://www.w3.org/TR/n-quads/>

<sup>19</sup>RDF/XML - <http://www.w3.org/TR/REC-rdf-syntax>

<sup>20</sup>Turtle - <http://www.w3.org/TR/turtle/>

<sup>21</sup>Trig - <http://www.w3.org/TR/trig/>

<sup>22</sup>RDFa - <http://www.w3.org/TR/rdfa-syntax/>

<sup>23</sup>(X)Html - <http://www.w3.org/TR/xhtml1/>

<sup>24</sup>JSON-LD - <http://www.w3.org/TR/json-ld/>

<sup>25</sup>Trojice nejsou z důvodu přehlednosti trojice uvedeny na samostatných řádcích

```

9      <http://rsmluv.cz/party/420> .
10
11 <http://rsmluv.cz/party/420>
12     <http://www.w3.org/1999/02/22-rdf-syntax-ns#type>
13     <http://purl.org/goodrelations/v1#BusinessEntity> .
14 <http://rsmluv.cz/party/420>
15     <http://purl.org/goodrelations/v1#legalName>
16     "Magistrát HMP" .
17 <http://rsmluv.cz/party/420>
18     <http://schema.org/address>
19     <http://rsmluv.cz/party/420/address> .
20
21 <http://rsmluv.cz/party/420/address>
22     <http://www.w3.org/1999/02/22-rdf-syntax-ns#type>
23     <http://schema.org/PostalAddress> .
24 <http://rsmluv.cz/party/420/address>
25     <http://schema.org/streetAddress>
26     "Staroměstské nám. 4" .

```

Výpis kódu 2.1: Příklad RDF dat - N-Triples

## 2.9.2 Příklad dat serializovaných ve formátu Turtle

Formát Turtle umožňuje zkracování URI pomocí prefixů. Umožňuje také zkracovat zápis tím, že nemusíme zapisovat opakující se subjekt. Jednotlivé dvojice predikát-hodnota lze tak přehledně mít u jednoho subjektu. Oddělovačem mezi dvojicemi v rámci subjektu je středník, blok informací o daném subjektu je zakončený tečkou. Pro definování typu subjektu se může použít klíčové slovo „a“, namísto predikátu `rdf:type`. Výhodou je úspornost a velmi dobrá lidská čitelnost (viz Kód 2.2).

```

1 @prefix cn: <http://tiny.cc/open-contracting#>.
2 @prefix dc: <http://purl.org/dc/terms/>.
3 @prefix gr: <http://purl.org/goodrelations/v1#>.
4 @prefix s:  <http://schema.org/>.
5
6 <http://rsmluv/contract/42/1> a cn:Contract;
7   dc:title "Softwarová zakázka";
8   cn:party <http://rsmluv/party/420>.
9
10 <http://rsmluv/party/420> a gr:BusinessEntity;
11   gr:legalName "Magistrát HMP";
12   s:address <http://rsmluv/party/420/address>.
13
14 <http://rsmluv/party/420/address> a s:PostalAddress;
15   s:streetAddress "Staroměstské nám. 4".

```

Výpis kódu 2.2: Příklad RDF dat - Turtle

Díky dobré čitelnosti, se formát Turtle hojně používá pro zapisování ontologií. V kódu 2.3 vidíme znázorněnou jednoduchou ontologii. Popisuje 2 objekty. Prvním je třída `Contract` (typ `owl:Class`). Definuje, že se jedná o smlouvu, je podtřídou (`rdfs:subClassOf`) třídy `Document` a je definována v ontologii (`rdfs:DefinedBy`) `http://tiny.cc/open-contracting`. Je to serializovaný zápis ontologie z obr. 2.7. Druhým objektem je vlastnost `party` (typ `owl:ObjectProperty`).

V predikátu `rdfs:domain` je specifikováno, že vlastnost `party` může být použita u třech tříd a to `Contract`, `Order` nebo `Invoice`. Predikát `rdfs:range` znamená, že očekávaný přiřazený objekt je typu `gr:BusinessEntity`.

```
1 @prefix :      <http://tiny.cc/open-contracting#> .
2 @prefix dc:    <http://purl.org/dc/terms/> .
3 @prefix gr:    <http://purl.org/goodrelations/v1#> .
4 @prefix owl: <http://www.w3.org/2002/07/owl#> .
5 @prefix rdf:   <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
6 @prefix rdfs:  <http://www.w3.org/2000/01/rdf-schema#> .
7
8 :Contract a owl:Class ;
9   rdfs:label "Smlouva"@cs, "Contract"@en ;
10  rdfs:subClassOf :Document ;
11  rdfs:isDefinedBy <http://tiny.cc/open-contracting> .
12
13 :party a owl:ObjectProperty ;
14   rdfs:label "Smluvní strana"@cs, "Party"@en ;
15   rdfs:domain [ a rdfs:Class ; owl:unionOf ( :Contract :Order
16     :Invoice ) ] ;
17   rdfs:range gr:BusinessEntity ;
18   rdfs:isDefinedBy <http://tiny.cc/open-contracting> .
```

Výpis kódu 2.3: Příklad RDF Ontologie - Turtle

### 2.9.3 Příklad dat serializovaných ve formátu JSON-LD

JSON-LD je jedním z poměrně nových formátů pro serializaci RDF. Jednou z motivací k vzniku byla snaha využít hojně využívané JSON dokumenty v dnešních aplikacích a co možná nejefektivněji z nich vytvořit RDF data.

Uveďme si modelový příklad. V kódu 2.4 jsou ne-RDF data ve formátu JSON. Jsou validní vůči nějakému JSON Schématu a používají se v konkrétních aplikacích.

V kódu 2.5 máme stejná data v RDF podobě. Jak je vidět, jednotlivým objektům je přiřazen typ a URI. Použije se k tomu klíčových slov `@type`, resp. `@id`. K dokumentu je také přiložen kontext (klíčové slovo `@context`), kde se definuje mapování vlastností původního JSON dokumentu na RDF ontologie. Zachovává se tedy původní struktura JSON dokumentu. Kontext však nemusí být přímo součástí JSON-LD dokumentu, lze se na něj odkazovat.

Výsledkem tedy může být JSON-LD soubor (viz Kód 2.6). Jedná se tedy pouze o lehce rozšířený původní JSON dokument. Z tohoto důvodu bude pravděpodobně takový dokument nadále validní vůči JSON Schématu a použitelný ve stávajících aplikacích. Přináší však tu výhodu, že se zároveň jedná o RDF data.

```

1 {
2   "title": "Softwarová zakázka",
3   "party": {
4
5     "name": "Magistrát HMP",
6     "address": {
7
8       "streetAddress": "Staroměstské nám. 4"
9     }
10  }
11 }

```

Výpis kódu 2.4: Obyčejný JSON dokument

```

1 {
2   "@context": "http://tiny.cc/open-contracting-context",
3
4   "@id": "http://rsmluv.cz/contract/42/1",
5   "@type": "Contract",
6   "title": "Softwarová zakázka",
7   "party": {
8
9     "@id": "http://rsmluv.cz/party/420",
10    "@type": "Party",
11    "name": "Magistrát HMP",
12    "address": {
13
14      "@id": "http://rsmluv.cz/party/420/address",
15      "@type": "Address",
16      "streetAddress": "Staroměstské nám. 4"
17    }
18  }
19 }

```

Výpis kódu 2.5: Příklad RDF dat - JSON-LD s Contextem

```

1 {
2   "@context": {
3
4     "cn": "http://tiny.cc/open-contracting#",
5     "dc": "http://purl.org/dc/terms/",
6     "gr": "http://purl.org/goodrelations/v1#",
7     "s": "http://schema.org/",
8
9     "Contract": "cn:Contract",
10    "Party": "gr:BusinessEntity",
11    "Address": "s:PostalAddress",
12    "title": "dc:title",
13    "party": "cn:party",
14    "name": "gr:legalName",
15    "address": "s:address",
16    "streetAddress": "s:streetAddress"
17  },
18
19   "@id": "http://rsmluv.cz/contract/42/1",
20   "@type": "Contract",
21   "title": "Softwarová zakázka",
22   "party": {
23

```



```

24  " @id" : "http://rsmluv.cz/party/420" ,
25  "@type" : "Party" ,
26  "name" : "Magistrát HMP" ,
27  "address" : {
28
29      " @id" : "http://rsmluv.cz/party/420/address" ,
30      "@type" : "Address" ,
31      "streetAddress" : "Staroměstské nám. 4"
32  }
33 }
34 }

```

Výpis kódu 2.6: Příklad RDF dat - JSON-LD

## 3. Otevřené smlouvy

### 3.1 Situace ve veřejné správě ČR

Pokud se veřejná instituce rozhodne pro publikaci údajů o smlouvách, má dnes (rok 2015) v podstatě dvě možnosti. První možností je vyvinutí vlastní iniciativy a zveřejnění smluv na svých webových stránkách. Druhou variantou je využití již existujícího registru smluv na portálu veřejné správy<sup>1</sup>. Registr je to značně minimalistický, ale řešení je to dostačující.

Vzhledem k chystanému zákonu o registru smluv<sup>2</sup> se ale budoucnost stávajícího registru jeví jako značně nejistá. Lze totiž očekávat, že s velkou pravděpodobností vznikne registr zbrusu nový.

První otázkou je, kolik veřejných institucí již smlouvy zveřejňuje. Na portálu veřejné správy lze dohledat řádově několik desítek subjektů. O těchto institucích můžeme prohlásit, že oficiálně zveřejňují smlouvy. Informace o subjektech, které zveřejňují na svých webových stránkách, není systematicky zdokumentovaná vůbec. Lze ale očekávat, vzhledem k celkovému množství veřejných institucí a počtu subjektů zveřejňujících na portálu veřejné správy, že se jedná o nepatrný zlomek. Klíčem ke zlepšení situace by mohl být již zmíněný zákon o registru smluv, který mimo jiné ukládá povinnost, že pokud smlouva není zveřejněná na internetu, tak je neplatná.

Další otázkou je, jak mají data o zveřejněných smlouvách vypadat, které položky musí, či nemusí obsahovat. Není přeci cílem, aby každá veřejná instituce zveřejňovala smlouvy jinak. Obecně chybí datový standard a metodika pro zveřejňování smluv. Pokrok v tomto směru udělalo Ministerstvo vnitra ČR, které na podzim roku 2015 plánuje vydat sadu standardů pro publikovatelné datové sady veřejných institucí<sup>3</sup>. Bude se mimo jiné jednat o jakési minimální nutné doporučení, co konkrétní datová sada musí obsahovat.

V úvodu již bylo řečeno, že pod záštitou Oživení o.s. a Centra aplikované ekonomie o.s. vzniká datový standard pro otevřené smlouvy. Hlavními postavami koordinujícími vývoj standardu se stali PhDr. Ing. Jiří Skuhrovec a Mgr. Lenka Franková. Na tvorbě standardu participují a mohou konstatovat, že základní verze je již hotová. Velmi pozitivní zprávou je to, že se tento standard s velkou pravděpodobností dostane do oficiálního doporučení Ministerstva vnitra ČR. Zdá se tedy, že celá tato snaha má smysl.

Standardem pro smlouvy to ale nekončí. Myšlenka úzké spolupráce zástupců měst a obcí, akademické a neziskové sféry se osvědčila. Výsledkem je vznik organizace Otevřená města<sup>4</sup>, která má za cíl sdružovat veřejné instituce. Pod společnou taktovkou pak financovat společné otevřené projekty. Prvním společným projektem je právě registr smluv.

---

<sup>1</sup>Portál veřejné správy - <http://portal.gov.cz>

<sup>2</sup>Návrh zákona o Registru smluv - tisk 42 - <http://www.psp.cz/sqw/historie.sqw?o=7&T=42>

<sup>3</sup>Standardy publikace a katalogizace otevřených dat veřejné správy ČR - <http://opendata.gov.cz/>

<sup>4</sup>Otevřená města - <http://www.otevrenamesta.cz/>

## 3.2 Standard pro zveřejňování smluv

V této kapitole se podrobněji seznámíme se standardem pro zveřejňování smluv. Nejdříve je vyložena základní struktura datového standardu, poté jsou popsány konkrétní položky standardu a číselníky. Následně jsou popsány způsoby publikace. Na závěr zmíníme několik informací o vznikající metodice pro zveřejňování smluv.

### 3.2.1 Základní struktura

“Základním objektem, který slouží k reprezentaci dat, je dokument. Jedná se o abstraktní entitu, která nabývá třech rozšířeních typu smlouva/příloha/dodatek. Tato rozšíření obsahují všechny položky obsažené v dokumentu a navíc konkrétní položky pro daný typ. Smluvní strany jsou separátní objekty navázané buď na smlouvu, objednávku nebo fakturu pomocí jednoznačného identifikátoru. Objedávka a faktura jsou separátní objekty, které se mohou vázat na konkrétní smlouvu/přílohu/dodatek pomocí jednoznačného identifikátoru. Rozšiřující entity mohou být součástí smlouvy, příp. objednávky. Reprezentují důležité události v životním cyklu dokumentu a jednotlivé transakce.

#### Reprezentované entity

- **Dokument** - základní abstraktní struktura pro evidování údajů o smlouvách/přílohách/dodatcích
  - **Smlouva** - detailní popisné údaje smlouvy
  - **Příloha** - popisné údaje přílohy
  - **Dodatek** - popisné údaje dodatku
  - **Vydavatel** - informace o vydavateli, který zveřejňuje údaje o smlouvách
  - **Verze** - identifikace jednotlivé verze dokumentu
- **Smluvní strana** - popisné údaje smluvní strany
  - **Nadřazená instituce** - informace o řídicí nebo ovládací právní osobě vystupující u smluvní strany
  - **Adresa** - podrobné údaje o adrese u smluvní strany
- **Objedávka** - popisné údaje objednávky, jedná se o doplňující informace k smlouvě/příloze/dodatku
- **Faktura** - popisné údaje faktury, jedná se o doplňující informace k smlouvě/příloze/dodatku
- **Rozšiřující entity** - rozšířené informace ke smlouvě, příp. objednávce
  - **Milník** - reprezentuje důležitou událost v životním cyklu smlouvy
  - **Transakce** - reprezentuje proběhlou platbu na základě smlouvy

Název pole	Popis
Název pole	Jméno reprezentující danou položku
Datový typ	Přípustný datový typ položky
Validita	Stupeň kvality položky.
Popis	Podrobný popis položky

Tabulka 3.1: zdroj: <http://standard.zindex.cz/doku.php/cs/standard/schema>

Datový model je rozdělen do tabulek podle jednotlivých reprezentovaných entit. Každá dílčí položka entity obsahuje tyto informace:

U každé zveřejněné smlouvy rozlišujeme tři stupně validity, resp. správnost a úplnost dat: A (kvalitní), B (dobrý), C (základní). Dokumenty musí splňovat alespoň minimální přípustnou kvalitu C. Pokud je nějaký atribut požadován pro stupeň validity C, je níže v textu označen např. takto (C). Položky doplněné systémem jsou označeny (S). Nepovinné položky jsou značeny (N), hvězdička znamená, že položka je kontrolována pokročilejším pravidlem popsáním u konkrétní položky.

Status	Validita	Popis
Nepovinné	N	Nepovinná položka
Základní	C	Povinná položka
Dobrý	B	Rozšiřující položka pro status „Dobrý“
Kvalitní	A	Rozšiřující položka pro status „Kvalitní“
Systémové	S	Položka doplněná systémem

Tabulka 3.2: Validita, zdroj: <http://standard.zindex.cz/doku.php/cs/standard/validate>

### Doplňující validační pravidla

Na entity se vztahují další validační pravidla, která nelze přehledně zachytit v rámci popisu jednotlivých položek. Jejich výčet je zde.

- Dokument je buď v strojově čitelném formátu (viz. Akceptovatelné soubory), nebo je k němu poskytnut plain text. Pro smlouvy účinné od 1.6.2015<sup>5</sup> je přípustná pouze varianta ve strojově čitelném formátu.
- U smlouvy typu darovací nesmí být připojeny faktury, ani jedna smluvní strana nesmí být identifikována jako Payer.
- Entita (Vydavatel/Smluvní strana/Nadřazená instituce) má vyplněno buď ID a nebo NoID = „true“.

<sup>5</sup>Předběžně, bude upřesněno

## Akceptovatelné soubory

Dokumenty připojené ke smlouvám by měly být strojově čitelné, resp. v těchto formátech:

Formát	Validita	Popis
PDF	C	Portable Document Format - ideálně strojově čitelný
DOC	C	Textový dokument Microsoft Word
XLS	C	Tabulka Microsoft Excel
DOCX	B	Textový dokument Microsoft Word
ODT	B	Textový dokument OpenDocument
XLSX	B	Tabulka Microsoft Excel
ODS	B	Tabulka OpenDocument

Tabulka 3.3: Akceptovatelné soubory,  
zdroj: <http://standard.zindex.cz/doku.php/cs/standard/validate>

### 3.2.2 Reprezentované entity

#### Dokument

Název pole	Datový typ	Validita	Popis
URI	String URI	S	Jednoznačný identifikátor formou URL. Typicky rsmluv.cz/[Typ]/[Id]/[Version], kde Version je vzestupné číslování verzí při změnách dokumentu či metadat
Document	String URI	S	Adresa URL fyzického umístění dokumentu. Typicky rsmluv.cz/[Typ]/[Id]/[Version]/File, viz akceptovatelné soubory
Versions	Object array	S	Údaje o verzi dokumentu. Viz entitia Verze
Type	String/ String enum	C	Typ dokumentu. Nabývá hodnot - Smlouva/Příloha/Dodatek
Publisher	Reference	C	Informace o vydavateli. Viz entitia Vydavatel
Valid	Boolean	B/S	Indikuje, zda dokument je platný, tj. nebyl zneplatněn nebo nahrazen novou verzí
PlainText	String	B/S	Prostý text dokumentu (nestrukturovaný, indexovatelný), alternativa pro scanované dokumenty
ResponsiblePersons	String array	B	Výčet odpovědných osob
Anonymised	Boolean	B	Značí, zda-li byla provedena anonymizace dokumentu

Tabulka 3.4: Vlastnosti dokumentu, zdroj:  
<http://standard.zindex.cz/doku.php/cs/standard/schema/contract>

#### Vydavatel

Název pole	Datový typ	Validita	Popis
ID	String	N	Identifikační číslo osoby, lze vložit i zahraniční ID
Name	String	C	Název, případně jméno a příjmení (s tituly)
NoID	Boolean	B	Indikuje že subjekt nemá IČO, nebo zahraniční ID
Country	String	B	Země původu, 3-písmený ISO kód
Authentication	String	S	Značí stupeň ověření zveřejňující strany

Tabulka 3.5: Vlastnosti vydavatele, zdroj:  
<http://standard.zindex.cz/doku.php/cs/standard/schema/contract>

## Verze

Název pole	Datový typ	Validita	Popis
PublisherId	String	N	Libovolný číselný identifikátor verze, spisové číslo apod.
Version	Int	S	Pořadové číslo verze, nejvyšší = aktuální
URI	String URI	S	Identifikátor dané verze
Published	DateTime	S	Datum publikace v systému

Tabulka 3.6: Vlastnosti verze smlouvy, zdroj:  
<http://standard.zindex.cz/doku.php/cs/standard/schema/contract>

## Smlouva

Název pole	Datový typ	Validita	Popis
AwardID	String	N*	Evidenční číslo veřejné zakázky. Uvádí se volitelně, pokud existuje
AwardProfileID	String	N	Číslo zakázky na profilu zadavatele
Amount	Nullable float	C*	Cena s DPH (u neplátců celková cena). Nejvyšší přípustná hodnota řádného plnění z dané smlouvy, které vynaloží některá smluvní strana. U smluv na dobu určitou se jedná o očekávané celkové finanční plnění strany s nejvyšším plněním, včetně opcí, bez sankcí. U smluv na dobu neurčitou, ve kterých není stanoven strop na celkové plnění, se jedná o nejvyšší očekávané roční plnění. U smluv bez finančního plnění (bartery, darovací smlouvy) je uvedena celková hodnota nefinančního plnění strany s nejvyšším plněním (např. odhadovaná hodnota daru). U smluv s nejasným plněním připustit NULL. Pokud je cena nenulová, tak alespoň jedna Smluvní strana (Party) musí mít příznak Payer = true
AmountNoVat	Nullable float	C*	Cena bez dph, uvádí se povinně pouze v případě, že Amount je s DPH
Title	String	C	Předmět smlouvy
ContractType	String	C	Číselník typů smlouvy, viz Číselníky
Parties	StringURI/Int array	C	Seznam identifikátorů (URI nebo LocalID) smluvních stran. Viz entitia Smluvní strana
SubjectType	String	B	Číselník typů zboží/služeb, viz. Číselníky
PriceAnnual	Boolean	B	Identifikuje, pokud je v Amount roční částka
Currency	String	B	Měna, 3-písmenný, ISO 4217 formát
DateSigned	Date	B	Datum posledního podpisu

Název pole	Datový typ	Validita	Popis
ValidFrom	Date	B	Datum účinnosti smlouvy
ValidUntil	Date	B	Datum ukončení účinnosti smlouvy (poslední plnění), NULL pro smlouvy na dobu
Funding	String	B	Převažující financování – vlastní, případně název dotačního titulu (bude kontrolován proti číselníku, viz. Číselníky)
Attachments	String URI-array	B	Seznam URI identifikátorů příloh. Viz entita Příloha
Amendments	String URI-array	B	Seznam URI identifikátorů dodatků. Viz entita Dodatek
Competency	String/String enum	A	Indikuje, zda-li se jedná o soukromoprávní nebo veřejnoprávní smlouvu
CurrentValidContract	String URI	A	Aktuálně platné znění smlouvy (se zapracovanými dodatky)
Description	String	A	Popis předmětu smlouvy
Implementation	Object	A	Objekt reprezentující transakce a milníky, viz entita Implementation

Tabulka 3.7: Vlastnosti smlouvy, zdroj:  
<http://standard.zindex.cz/doku.php/cs/standard/schema/document>

## Příloha

Název pole	Datový typ	Validita	Popis
Title	String	C	Název
Contract	String URI	C	Jednoznační identifikátor smlouvy
Number	Int	B	Číslo přílohy

Tabulka 3.8: Vlastnosti přílohy, zdroj:  
<http://standard.zindex.cz/doku.php/cs/standard/schema/attachment>

## Dodatek

Název pole	Datový typ	Validita	Popis
Title	String	C	Název
Contract	String URI	C	Jednoznační identifikátor smlouvy
Number	Int	B	Pořadové číslo dodatku (podle času podpisu)
DateSigned	Date	B	Datum podpisu

Tabulka 3.9: Vlastnosti dodatku, zdroj:  
<http://standard.zindex.cz/doku.php/cs/standard/schema/amendment>



## Smluvní strana

Název pole	Datový typ	Validita	Popis
ID	String	N	Identifikační číslo osoby, lze vložit i zahraniční id
LocalID	String URI/Int	C	Jednoznačný identifikátor v rámci dokumentu
Name	String	C	Název, případně jméno a příjmení (s tituly)
Payer	Boolean	C*	Identifikuje stranu která bude finančně plnit, pokud není zřejmé, nevyplňuje se
NoID	Boolean	B	Indikuje že subjekt nemá IČO, nebo zahraniční ID
Country	String	B	Země původu, 3-písmený ISO kód
Address	String/Referer	A	Adresa subjektu, případně "Anonymizováno". Umožňuje zadat adresu jako prostý řetězec, nebo strukturovaně, viz entitia Adresa
PaysVAT	Boolean	A	Indikuje, zda-li je subjekt plátce DPH
SuperiorInstitution	Reference	N/S	Řídící nebo ovládající právnická osoba, v případě veřejnoprávních smluv nadřazený správní orgán. Viz. Nadřazená instituce

Tabulka 3.10: Vlastnosti smluvní strany, zdroj:  
<http://standard.zindex.cz/doku.php/cs/standard/schema/contract>

## Nadřazené instituce

Název pole	Datový typ	Validita	Popis
ID	String	N	Identifikační číslo osoby, lze vložit i zahraniční id
LocalID	String URI/Int	C	Jednoznačný identifikátor v rámci dokumentu
Name	String	C	Název, případně jméno a příjmení (s tituly)
NoID	Boolean	B	Indikuje že subjekt nemá IČO, nebo zahraniční ID
Country	String	B	Země původu, 3-písmený ISO kód

Tabulka 3.11: Vlastnosti nadřazené instituce, zdroj:  
<http://standard.zindex.cz/doku.php/cs/standard/schema/contract>

## Adresa

Název pole	Datový typ	Validita	Popis
StreetAddress	String	A	Ulice, případně "Anonymizováno"
Locality	String	A	Město, případně "Anonymizováno"
PostalCode	Integer	A	PSČ, případně "Anonymizováno"
Nuts	String	A	Normalizovaná klasifikace územních celků (např. Praha - CZ010), případně "Anonymizováno"

Tabulka 3.12: Vlastnosti adresy, zdroj:  
<http://standard.zindex.cz/doku.php/cs/standard/schema/contract>

## Objednávka

Název pole	Datový typ	Validita	Popis
ParrentDocument	String URI	N	Jednoznačný identifikátor dokumentu
SubjectType	String	N	Číselník typů zboží/služeb, viz. Číselníky
Parties	String URI/Int array	N	Seznam identifikátorů (URI nebo LocalID) smluvních stran. Viz entitia Smluvní strana
Title	String	C	Předmět
Amount	Float	C	Cena s DPH
Currency	String	B	Měna, 3-písmenný, ISO 4217 formát
DateSigned	Date	B	Datum posledního podpisu
Implementation	Object	A	Objekt reprezentující transakce a milníky, viz entitia Implementation

Tabulka 3.13: Vlastnosti objednávky, zdroj:  
<http://standard.zindex.cz/doku.php/cs/standard/schema/order>

## Faktura

Název pole	Datový typ	Validita	Popis
ParrentDocument	String URI	N	Jednoznačný identifikátor dokumentu
Parties	String URI/Int array	N	Seznam identifikátorů (URI nebo LocalID) smluvních stran. Viz entitia Smluvní strana
Title	String	C	Předmět
Amount	Float	C*	Cena s DPH (u neplátců celková cena).
Currency	String	B	Měna, 3-písmenný, ISO 4217 formát
DateSigned	Date	B	Datum posledního podpisu

Název pole	Datový typ	Validita	Popis
DueDate	Date	B	Datum splatnosti

Tabulka 3.14: Vlastnosti faktury, zdroj:  
<http://standard.zindex.cz/doku.php/cs/standard/schema/invoice>

## Rozšiřující entity

### Implementace

Název pole	Datový typ	Validita	Popis
Milestones	Object arra	A	Milníky, pro volnou evidenci událostí (obnova smlouvy, předání apod.). Viz entitia Milník
Transactions	Object array	A	Seznam transakcí, tedy proběhlých plateb na základě smlouvy. Viz entitia Transakce

Tabulka 3.15: Vlastnosti implementace, zdroj:  
<http://standard.zindex.cz/doku.php/cs/standard/schema/implementation>

## Milník

Název pole	Datový typ	Validita	Popis
Title	String	C	Název
DueDate	String	C	Datum

Tabulka 3.16: Vlastnosti milníku, zdroj:  
<http://standard.zindex.cz/doku.php/cs/standard/schema/implementation>

## Transakce

Název pole	Datový typ	Validita	Popis
Date	DateTime	C	Datum a čas proběhlé transakce
Ammount	Float	C	Zaplacená cena s DPH, vždy stejná měna jako v Currency
SenderOrganization	Reference	C	Informace o odesílateli. Viz entitia Party
ReceiverOrganization	Reference	C	Informace o příjemci. Viz entitia Party
PublisherId	String	B	Libovolný číselný identifikátor transakce, unikátní v rámci smlouvy

Tabulka 3.17: Vlastnosti transakce, zdroj:  
<http://standard.zindex.cz/doku.php/cs/standard/schema/implementation>

### 3.2.3 Číselníky

V následujících tabulkách jsou znázorněny přípustné hodnoty číselníků Typ dokumentu (vlastnost Type u entity Dokument) a Typ smlouvy (vlastnost ContractType u entity Smlouva). Číselník Typ zboží a služeb (položka SubjectType u entity Smlouva) je zveřejněn na portálu informačního systému o veřejných zakázkách<sup>6</sup>.

Hodnota
Smlouva
Příloha
Dodatek

Tabulka 3.18: Číselník typu dokumentu, zdroj:  
<http://standard.zindex.cz/doku.php/cs/standard/codelists>

Hodnota
Nájemní smlouva
Darovací smlouva
Kupní smlouva
Směnná smlouva
Pojistná smlouva
Smlouva o výpůjčce
Licenční smlouva
Mandátní smlouva
Leasingová smlouva
Pachtovní smlouva
Smlouva o zřízení věcného břemene
Smlouva o provedení stavby
Smlouva o provedení práce
Smlouva o provedení uměleckého výkonu
Smlouva o úvěru
Smlouva o uzavření budoucí smlouvy
Veřejnoprávní smlouva
Jiná

Tabulka 3.19: Číselník typu smlouvy, zdroj:  
<http://standard.zindex.cz/doku.php/cs/standard/codelists>

---

<sup>6</sup>Číselníky a klasifikace na portálu informačního systému o veřejných zakázkách - TODO

## 3.3 Publikace

Pro potřeby publikace je třeba zvolit vhodný datový formát v kterém budou otevřené smlouvy přenositelné. Jako kritéria výběru vhodného formátu stanovíme čtyři podmínky:

- otevřený datový formát - tím zaručíme otevřená data na úrovni kvality 3★
- obecná znalost a jednoduchost datového formátu - cílem je, aby valná většina IT specialistů ve veřejných institucích formát znala
- existence volně dostupných nástrojů k čtení a zpracování datového formátu
- možnost tvorby datového schématu - resp. možnost určit soustavu speciﬁkáci a pravidel, jak má datový soubor vypadat, aby byl validní

Není překvapující, že obecně nejznámějšími datovými formáty splňujícími výše zmíněná pravidla jsou formáty XML (Extensible Markup Language) a JSON (JavaScript Object Notation)<sup>7</sup>. Vzhledem k úspornosti a možností rychlejšího zpracování padla volba na formát JSON.

Pokud však chceme, aby datový standard byl součástí plánovaného doporučení Ministerstva vnitra ČR, tak je nutné podporovat také formát CSV (Comma-separated values)<sup>8</sup>. Jedná se o jednoduchý, otevřený datový formát, ale s plochou strukturou. Publikace smluv v CSV si tedy vyžádá řadu omezení.

### 3.3.1 JSON

Základní strukturu datového souboru lze vidět z Tabulky 3.20. Položky Id, Date a Language slouží k popisu datového souboru jako celku. Položky Documents, Parties, Orders a Invoices už obsahují konkrétní výčty entit ze standardu. Položky vyznačené stupněm validity C, jsou povinné.

Ke konkrétní speciﬁkaci jednotlivých položek ve formátu JSON se používá JSON Schema<sup>9</sup>. Lze v něm definovat konkrétní elementy a podelementy, výchozí hodnoty, datové typy, požadovaný obsah apod. Příklad JSON Schématu vycházejícího z datového standardu lze nalézt v příloze A.

Datový soubor, validní vůči JSON schématu, s jednou smlouvou a dvěma smluvními stranami můžeme vidět na příkladu kódu 3.1.

---

<sup>7</sup>Datový formát JSON - <https://www.ietf.org/rfc/rfc4627.txt>

<sup>8</sup>Datový formát CSV - <https://tools.ietf.org/html/rfc4180>

<sup>9</sup>JSON Schema - <http://json-schema.org/documentation.html>, popis způsobu zápisu konkrétních položek je nad rámce této práce

```

1 {
2   "id": "89f689cd-e784-4374-bb17-94144679d46f",
3   "published": "2014-03-25T23:20:50+01:00",
4   "language": "cs",
5
6   "documents": [
7     {
8       "uri": "http://rsmluv.cz/smlouva/12345",
9       "document": "http://rsmluv.cz/smlouva/12345/Smlouva12345.docx",
10      ,
11      "type": "Smlouva",
12      "valid": true,
13      "anonymised": false,
14
15      "awardID": "486026",
16      "awardProfileID": "OI-010143",
17      "amount": 584520.00,
18      "title": "Brno, Vackova, Šafaříkova – rekonstrukce kanalizace
19      a vodovodu",
20      "contractType": "Kupní smlouva",
21      "subjectType": "Právní, finanční překladatelské, pojišťovnické
22      , poradenské a jiné služby",
23      "priceAnnual": false,
24      "currency": "CZK",
25      "dateSigned": "2011-11-16",
26      "validFrom": "2011-11-02",
27      "validUntil": "2012-06-30",
28      "funding": "vlastní",
29      "competency": [ "Soukromoprávní smlouva" ],
30      "currentValidContract": "http://zakazky.brno.cz/?pg=detail&id
31      =18249&list=135",
32      "description": "Projektová dokumentace pro stavební povolení a
33      zadání stavby bude řešit rekonstrukci stávající kanalizační
34      stoky z profilu DN 500 na DN 800/1200 v délce 146 m, rekonstrukci
35      kanalizačních přípojek pod veřejným prostranstvím a přepojení vš
36      ech stávajících dešťových vpustí a také vybourání vozovek a chodn
37      íků nad rýhou a zásyp rýhy recyklátem. Součástí bude inženýrsko-
38      geologický průzkum, geodetické zaměření dotčené oblasti,
39      inventarizace zeleně, výkaz výměr, položkový rozpočet a výkon
40      autorského dozoru až do dokončení stavby. Dokumentace bude
41      projednána s orgány státní správy a s účastníky stavebního řízení
42      a jejich připomínky budou do dokumentace zapracovány.",
43
44      "responsiblePersons": [ "Ing. Petr Vokřál ", "Mgr. Adriana
45      Krnáčová, MBA" ],
46
47      "publisher": {
48        "id": "6003508",
49        "name": "Statutární město Brno",
50        "noID": false,
51        "country": "CZE",
52        "authentication": "email"
53      },
54
55      "parties": [ 132456, 987654 ],
56
57      "Implementation": {
58        "milestones": [

```

```

44         {
45             "title" : "Výpověď smlouvy",
46             "dueDate" : "2012-06-20T23:20:50+01:00"
47         }
48     ],
49     "transactions" : [
50         {
51             "publisherId" : "1269483",
52             "date" : "2012-01-01T18:35:20+01:00",
53             "amount" : 300000,
54             "senderOrganization" : 987654,
55             "receiverOrganization" : 132456
56         },
57         {
58             "publisherId" : "934584",
59             "date" : "2012-02-01T09:13:40+01:00",
60             "amount" : 284520,
61             "senderOrganization" : 987654,
62             "receiverOrganization" : 132456
63         }
64     ]
65 },
66
67 "versions" : [
68     {
69         "version" : 1,
70         "uri" : "http://rsmluv.cz/smlouva/12345/verze/1",
71         "published" : "2014-09-15T23:20:50+01:00"
72     },
73     {
74         "version" : 2,
75         "uri" : "http://rsmluv.cz/smlouva/12345/verze/2",
76         "published" : "2015-03-15T14:35:28+01:00"
77     }
78 ]
79 }
80 ],
81
82 "parties" : [
83     {
84         "id" : "44992785",
85         "localID" : 132456,
86         "name" : "Statutární město Brno",
87         "payer" : false,
88         "noID" : false,
89         "country" : "CZE",
90         "address" : {
91             "streetAddress" : "Dominikánské náměstí 196/1",
92             "locality" : "Brno-město, Brno",
93             "postalCode" : 60200,
94             "nuts" : "CZ064"
95         },
96         "superiorInstitution" : {
97             "id" : "00064581",
98             "localID" : 56486,
99             "name" : "Magistrát hlavního města Prahy",
100             "noID" : false,
101             "country" : "CZE"

```

```

102     }
103   },
104   {
105     "id" : "46347011",
106     "localID" : 987654,
107     "name" : "Kovoprojekta Brno a.s.",
108     "payer" : true,
109     "noID" : false,
110     "country" : "CZE",
111     "address" : {
112       "streetAddress" : "Šumavská 416/15",
113       "locality" : "Ponava, Brno",
114       "postalCode" : 60200,
115       "nuts" : "CZ064"
116     }
117   }
118 ]
119 }

```

Výpis kódu 3.1: JSON soubor s jednou smlouvou



### 3.3.2 CSV

Formát CSV je jednoduchou plochou strukturou, nelze tedy pomocí tohoto formátu zaznamenat úplnou strukturu datového standardu. Řešením by mohlo být rozdělit údaje o smlouvách do sady CSV souborů. Tím se ale ztrácí výhoda jednoduchosti CSV. Cílem publikace v CSV je maximální jednoduchost pro vydavatele. Proto jsme přistoupili k následujícím omezením<sup>10</sup>

- Vše je smlouvou, tedy nebudeme evidovat dodatky, přílohy, faktury a objednávky
- Vypuštěny/omezeny vlastnosti u smlouvy
  - URI - nahrazeno odkazem na podrobné údaje o smlouvě s validitou A
  - Type
  - Verzování, resp. vypuštěna vazba na verze a vlastnost Valid
  - PlainText
  - Vydavatel, převážně proto, že MV má pro vydavatele speciální strukturu
  - Pouze jedna zodpovědná osoba
  - Vypuštěny rozšiřující entity - milníky a transakce
- Umožněny pouze dvě smluvní strany - Publisher a Partner, a to s vlastnostmi
  - ID, Name, Payer, Country, Address a PaysVAT
- Nové vlastnosti
  - LocalID - Libovolný číselný identifikátor smlouvy, spisové číslo apod. - validita A
  - NumberOfAmendments - Počet dodatků - validita B
  - LastAmendmentDateSigned - Datum podpisu posledního dodatku - validita A
  - FirstInvoiceDueDate - Datum splatnosti první faktury - validita A
  - LastInvoiceDueDate - Datum splatnosti poslední faktury - validita A
  - TotalFillingValue - Celkový objem plnění - validita A

---

<sup>10</sup>Stav k červenci 2015

### 3.4 Metodika zveřejňování smluv

Spolu s datovým standardem vzniká i metodika mající za cíl technicky i věcně datový standard popsat. Tvorbu této metodiky jsem pod taktovkou Jiřího Skuhrovce dostal na starost.

Jedná se o jednoduchou webovou aplikaci na bázi wikipedie<sup>11</sup>. Implementována je pomocí nástroje Dokuwiki<sup>12</sup>.

---

<sup>11</sup>Wikipedia TODO

<sup>12</sup>Dokuwiki TODO

## 4. Otevřené smlouvy jako Linked Data

V minulé kapitole bylo řečeno, jak reprezentovat smlouvy jako otevřená data. Data serializovaná do formátu JSON, či CSV můžeme kategorizovat stupněm 3★[TODO - odkaz na kapitolu]. Pokud však chceme dosáhnout otevřenosti dat kategorie 5★, je třeba provést několik dalších kroků:

- Identifikovat reprezentované objekty a vlastnosti pomocí URI
- Vytvořit strojově srozumitelné struktury, resp. napojit data na konkrétní slovníky tříd a predikátů - ontologie
- Propojit smlouvy pomocí odkazů na související data

### 4.1 Přiřazení identifikátorů jednotlivým entitám otevřených smluv

K jednoznačné identifikaci každé entity nám stačí její *typ* a *Id.* Výjimku tvoří dokumenty, které jsou verzované. K nim je nutné přidat informaci o konkrétní verzi. Dále chceme vyjádřit vztah podřízené entity k nadřízené. Řešením je opět přidání informací o typu podřízené entity, příp. jejího *Id.* Resp. základní URI schéma bude:

*http://[domain]/[typ]/[entitaId]/[versionId]/[childEntity]/[childEntityId]*

- domain je doména instituce publikující smlouvy
- údaje ve složených závorkách jsou volitelné

Výsledné identifikátory entit:

- **Document** (Dokument) - *http://[domain]/[type]/[documentId]/[versionId]*  
Type může nabývat hodnot contract/attachment/amendment - resp. jedná se v podstatě o hodnotu položky Uri z datového standardu
- **Publisher TODO** (Vydavatel) -  
*http://[domain]/[type]/[documentId]/[versionId]/publisher* Jedná se o podřízenou položku dokumentu
  - Jedná se o podřízenou položku dokumentu
- **Version** (Verze) - *http://[domain]/[type]/[documentId]/[versionId]/version*  
Jedná se o podřízenou položku dokumentu
  - Jedná se o podřízenou položku dokumentu
- **Order** (Objednávka) - *http://[domain]/order/[orderId]*

- **Invoice** (Faktura) - *http://[domain]/invoice/[invoiceId]*
- **Implementation** (Implementace)
  - *http://[domain]/[type]/[documentId]/[versionId]/implementation*
    - \* Pokud se jedná o podřízenou položku dokumentu
  - *http://[domain]/order/[orderId]/implementation*
    - \* Pokud se jedná o podřízenou položku objednávky
- **Milestone** (Milník)
  - Zde si dovolíme drobné zjednodušení, a to vynechání implementačního z identifikátoru. Adresa bude jednodušší, informační hodnota však zůstane stejná
  - *http://[domain]/[type]/[id]/[versionId]/milestone/[milestoneId]*
    - \* Milník u dokumentu
  - *http://[domain]/order/[orderId]/milestone/[milestoneId]*
    - \* Milník u objednávky
- **Transaction** (Transakce)
  - Zjednodušení, viz Milník
  - *http://[domain]/[type]/[id]/[versionId]/transaction/[transactionId]*
    - \* Transakce u dokumentu
  - *http://[domain]/order/[orderId]/transaction/[transactionId]*
    - \* Transakce u objednávky
- **Party** (Smluvní strana) - *http://[domain]/party/[partyId]*
- **Address** (Adresa) - *http://[domain]/party/[partyId]/address*
  - Jedná se o podřízenou položku smluvní strany
- **SuperiorInstitution** (Nadřazená instituce) - *http://[domain]/superiorInstitution/[superiorInstitutionId]*

## 4.2 Ontologie pro publikaci dat o smlouvách

Než začneme s tvorbou ontologie, je dobré si uvědomit, že vycházíme z již hotového datového standardu. Nemáme tedy v tvorbě úplnou svobodu. Cílem tedy bude tvorba takové ontologie, která bude odpovídat stávajícímu datovému standardu, a přesto se bude snažit využít co nejvíce již existujících ontologií.

Samotnou tvorbu ontologie rozdělíme do dvou kroků:

- Analýza vhodných, již existujících ontologií
- Tvorba samotné ontologie

### 4.2.1 Analýza vhodných, již existujících ontologií

Při tvorbě ontologie se zaměříme na otázku, zdali existuje třída, či predikát v nějaké ontologii sémanticky ekvivalentní třídě, či konkrétní položce datového standardu pro smlouvy<sup>1</sup>. Takových vhodných ontologií ve světě Linked Data může být celá řada. K výběru stačí libovolná z nich.

V následujícím seznamu je výčet vybraných ontologií, které se jeví jako vhodné pro použití při popisování smluv<sup>2</sup>. U každého bodu je zmíněn popis ontologie, důvod, proč byla daná ontologie zvolena, a seznam tříd a predikátů vhodných k použití.

Vybrané ontologie:

- **Commerce** (<https://w3id.org/commerce#>) - ontologie pro popisování obchodních transakcí
  - Důvod použití - užitečná třída transakce
  - Vybrané třídy
    - \* Transaction - třída reprezentující transakci
  - Vybrané predikáty
    - \* contentUrl - adresa obsahu
    - \* source - zdroj transakce (pozor na party)
    - \* destination - cíl transakce (pozor na party)
- **DublinCore** (<http://purl.org/dc/terms/>) - základní ontologie pro popis metadat
  - Důvod použití - základní a všeobecně známá ontologie popisující metadata
  - Vybrané predikáty
    - \* created - datum vytvoření
    - \* creator - tvůrce
    - \* date - obecné datum
    - \* description - popis metadat
    - \* identifier - jednoznačný identifikátor
    - \* issued - datum publikace
    - \* language - jazyk
    - \* modified - datum modifikace
    - \* publisher - vydavatel
    - \* rights - licence
    - \* title - název dokumentu
    - \* type - typ dokumentu

---

<sup>1</sup>TODO

<sup>2</sup>Obecně výběr ontologií nemusíme považovat za striktní. Každou třídu, či predikát lze označit jako sémanticky ekvivalentní jiné třídě, či predikátu. Slouží k tomu konstrukce jazyka owl[TODO - odkaz definice ontologie] - *equivalentClass*, resp. *equivalentProperty*.

- **Friend-of-a-Friend** (<http://xmlns.com/foaf/0.1>) - ontologie pro popis vazez mezi lidmi
  - Důvod použití - vhodná pro označení třídy vydavatele
  - Vybrané třídy
    - \* Person - třída reprezentující osobu
    - \* Organization - třída reprezentující organizaci
  - Vybrané predikáty
    - \* name - jméno osoby
    - \* mbox - email osoby
- **GoodRelations** (<http://purl.org/goodrelations/v1#>) - ontologie pro popis produktů, cen a obchodních dat
  - Důvod použití - známá ontologie, vhodná pro popis smluvních stran a informací o cenách
  - Vybrané třídy
    - \* BusinessEntity - třída popisující hospodářské subjekty
  - Vybrané predikáty
    - \* hasCurrency - měna
    - \* hasCurrencyValue - cena
    - \* legalName - název subjektu
    - \* valueAddedTaxIncluded - plátce DPH
- **PaySwarm** (<https://w3id.org/payswarm#>) - ontologie popisující účtenky, platby, pronájmy a obecně výměnné platby na webu
  - Důvod použití - obsahuje predikáty popisující intervaly platnosti
  - Vybrané predikáty
    - \* validFrom - datum platnosti od
    - \* validUntil - datum platnosti do
- **Schema** (<http://schema.org/>) - obecná ontologie mající za cíl pokrývat co největší možné množství informací
  - Důvod použití - známá ontologie, možnost využití pro popis adresních údajů
  - Vybrané třídy
    - \* PostalAddress - třída reprezentující adresu
  - Vybrané predikáty
    - \* address - adresa
    - \* addressCountry - země
    - \* addressLocality - město
    - \* postalCode - PSČ

\* streetAddress - ulice

- **Vann** (<http://purl.org/vocab/vann/>) - anotační ontologie pro dokumenty
  - Důvod použití - nesouvisí s datovým standardem, tato ontologie je vhodná pro popis ontologií a bude zmíněna níže v publikaci[TODO-odkaz].
  - Vybrané predikáty
    - \* preferredNamespaceUri - preferovaná adresa ontologie
    - \* preferredNamespacePrefix - preferovaná zkratka
    - \* usageNote - poznámka k použití

## 4.2.2 Tvorba ontologie

Každou položku datového standardu namapujeme na třídu, či predikát výsledné ontologie. Pro některé položky využijeme zmíněné predikáty z již existujících ontologií, pro ostatní vytvoříme třídy a predikáty vlastní.

Vlastní ontologii nazveme jako open-contracting a budeme pro ni používat zkratku cn.

Výsledné mapování můžeme vidět v následujících tabulkách 4.1 - 4.13. V prvním sloupečku se nachází entita/vlastnost datového standardu, v druhém namapovaná třída/predikát a v třetím případná poznámka<sup>3</sup>.

### Dokument

Entita/Vlastnost	Třída/Predikát	Poznámka
<b>uri</b>	cn:uri	
<b>document</b>	com:contentUrl	
<b>versions</b>	cn:version	
<b>type</b>	dcmi:type	
<b>publisher</b>	dc:publisher	
<b>valid</b>	cn:valid	
<b>plainText</b>	cn: plainText	
<b>responsiblePersons</b>	cn:responsiblePerson	
<b>anonymised</b>	cn:anonymised	

Tabulka 4.1: Mapování entity Document

---

<sup>3</sup>Entity jsou uváděny velkým písmem, predikáty malým

## Vydavatel

Entita/Vlastnost	Třída/Predikát	Poznámka
<b>Publisher</b>	foaf:Organization	
<b>id</b>	dc:identifier	
<b>name</b>	gr:legalName	
<b>noID</b>	cn:noID	
<b>country</b>	schema:addressCountry	
<b>authentication</b>	cn:authentication	

Tabulka 4.2: Mapování entity Vydavatel

## Verze

Entita/Vlastnost	Třída/Predikát	Poznámka
<b>Version</b>	cn:Version	
<b>publisherId</b>	cn:publisherId	
<b>version</b>	cn:versionOrder	
<b>uri</b>	cn:uri	
<b>published</b>	dc:issued	

Tabulka 4.3: Mapování entity Verze smlouvy

## Smlouva

Entita/Vlastnost	Třída/Predikát	Poznámka
<b>Contract</b>	cn:Contract	
<b>awardID</b>	cn:awardID	
<b>awardProfileID</b>	cn:awardProfileID	
<b>amount</b>	gr:hasCurrencyValue	
<b>amountNoVat</b>	gr:hasCurrencyValue	
<b>title</b>	dc:title	
<b>contractType</b>	cn:contractType	
<b>parties</b>	cn:party	
<b>subjectType</b>	cn:subjectType	
<b>priceAnnual</b>	cn:priceAnnual	
<b>currency</b>	gr:hasCurrency	
<b>dateSigned</b>	dc:created	
<b>validFrom</b>	ps:validFrom	
<b>validUntil</b>	ps:validUntil	



Entita/Vlastnost	Třída/Predikát	Poznámka
<b>funding</b>	cn:funding	
<b>attachments</b>	cn:attachment	
<b>amendments</b>	cn:amendment	
<b>competency</b>	cn:competency	
<b>currentValidContract</b>	cn:currentValidContract	
<b>description</b>	dc:description	
<b>implementation</b>	cn:implementation	

Tabulka 4.4: Mapování entity Smlouva

## Příloha

Entita/Vlastnost	Třída/Predikát	Poznámka
<b>Attachment</b>	cn:Attachment	
<b>title</b>	dc:title	
<b>contract</b>	cn:contract	
<b>number</b>	cn:attachmentOrder	

Tabulka 4.5: Mapování entity Příloha

## Dodatek

Entita/Vlastnost	Třída/Predikát	Poznámka
<b>Amendment</b>	cn:Amendment	
<b>title</b>	dc:title	
<b>contract</b>	cn:contract	
<b>number</b>	cn:amendmentOrder	
<b>dateSigned</b>	dc:created	

Tabulka 4.6: Mapování entity Dodatek

## Smluvní strana

Entita/Vlastnost	Třída/Predikát	Poznámka
<b>Party</b>	gr:BusinessEntity	
<b>id</b>	dc:identifier	
<b>localID</b>	cn:localID	
<b>name</b>	gr:legalName	
<b>payer</b>	cn:payer	

Entita/Vlastnost	Třída/Predikát	Poznámka
<b>noID</b>	cn:noID	
<b>country</b>	schema:addressCountry	
<b>paysVAT</b>	gr:valueAddedTaxIncluded	
<b>superiorInstitution</b>	cn:superiorInstitution	

Tabulka 4.7: Mapování entity Smluvní strana

### Nadřazená instituce

Entita/Vlastnost	Třída/Predikát	Poznámka
<b>SuperiorInstitution</b>	gr:BusinessEntity	
<b>id</b>	dc:identifier	
<b>localID</b>	cn:localID	
<b>name</b>	gr:legalName	
<b>payer</b>	cn:payer	
<b>noID</b>	cn:noID	
<b>country</b>	schema:addressCountry	

Tabulka 4.8: Mapování entity Nadřazená instituce

### Adresa

Entita/Vlastnost	Třída/Predikát	Poznámka
<b>Address</b>	schema:PostalAddress	
<b>streetAddress</b>	schema:streetAddress	
<b>locality</b>	schema:addressLocality	
<b>postalCode</b>	schema:postalCode	
<b>nuts</b>	cn:nuts	

Tabulka 4.9: Mapování entity Nadřazená instituce

### Objednávka

Entita/Vlastnost	Třída/Predikát	Poznámka
<b>Order</b>	cn:Order	
<b>parentDocument</b>	cn:parentDocument	
<b>subjectType</b>	cn:subjectType	
<b>parties</b>	cn:party	

Entita/Vlastnost	Třída/Predikát	Poznámka
<b>title</b>	dc:title	
<b>amount</b>	gr:hasCurrencyValue	
<b>currency</b>	gr:hasCurrency	
<b>dateSigned</b>	dc:created	
<b>implementation</b>	cn:implementation	

Tabulka 4.10: Mapování entity Objednávka

## Faktura

Entita/Vlastnost	Třída/Predikát	Poznámka
<b>Invoice</b>	cn:Invoice	
<b>parentDocument</b>	cn:parentDocument	
<b>parties</b>	cn:party	
<b>title</b>	dc:title	
<b>amount</b>	gr:hasCurrencyValue	
<b>currency</b>	gr:hasCurrency	
<b>dateSigned</b>	dc:created	
<b>dueDate</b>	cn:dueDate	

Tabulka 4.11: Mapování entity Faktura

## Rozšiřující entity

### Implementace

Entita/Vlastnost	Třída/Predikát	Poznámka
<b>Implementation</b>	cn:Implementation	
<b>milestones</b>	cn:milestone	
<b>transactions</b>	cn:transaction	

Tabulka 4.12: Mapování entity Implementace

### Milník

Entita/Vlastnost	Třída/Predikát	Poznámka
<b>Milestone</b>	cn:Milestone	
<b>title</b>	dc:title	
<b>dueDate</b>	cn:dueDate	

Tabulka 4.13: Mapování entity Milník

## Transakce

Entita/Vlastnost	Třída/Predikát	Poznámka
<b>Transaction</b>	com:Transaction	
<b>date</b>	dc:date	
<b>amount</b>	gr:hasCurrencyValue	
<b>senderOrganization</b>	com:source	
<b>receiverOrganization</b>	com:destination	
<b>publisherId</b>	cn:publisherId	

Tabulka 4.14: Mapování entity Transakce

### 4.2.3 Publikace

K serializaci výsledné ontologie využijeme formátu Turtle. Soubor se skládá z hlavičky a definicí nově vytvořených tříd a predikátů. V hlavičce definujeme prefixy použitých ontologií a základní informace o ontologii. Použité třídy a predikáty zmíníme v poznámkách k použití (predikát *vann:usageNote*). Příklad ontologie lze nalézt v Příloze B(TODO).

## 4.3 Možnosti propojení na související data

První bezpochyby zajímavou možností je propojení smluv s veřejnými zakázkami, resp. věstníkem veřejných zakázek provozovaným Ministerstvem pro místní rozvoj. Jednotlivé smlouvy mající spojitost s veřejnou zakázkou poznáme podle vlastnosti AwardID, resp. evidenční číslo veřejné zakázky.

Dalšími zajímavými prvky k propojení jsou pokročilé informace o smluvním stranách. Každá smluvní strana vystupující ve smlouvě může mít zveřejněno mimo jiné identifikační číslo a adresu. Nabízí se tedy propojení s národními registry ARES a RÚIAN. ARES je registrem informací o ekonomických subjektech provozovaný Ministerstvem financí, RÚIAN je registrem územní identifikace, adres a nemovitostí provozovaný Českým úřadem zeměměřickým a katastrálním. Vhodné informace k propojení tedy jsou:

- Evidenční číslo veřejné zakázky u smlouvy na věstník veřejných zakázek
  - Iniciativa OpenData.cz zpracovává údaje o veřejných zakázkách v RDF podobě, využijeme tedy propojení právě s tímto datasetem
  - Cílové URL - <http://linked.opendata.cz/resource/business-entity/CZICO>
- Identifikační číslo smluvní strany na ARES
  - OpenData.cz aktuálně zpracovávají také údaje z registru ARES
  - Cílové URL - <http://linked.opendata.cz/resource/domain/buyer-profiles/contract/cz>

- Adresa smluvní strany na RÚIAN - TODO

## 4.4 Provázání s datovým formátem JSON

V třetí kapitole jsme si ukázali, jak publikovat smlouvy v datovém formátu JSON. Vzhledem k budoucímu možnému využití v aplikacích pracujících nad smlouvami v JSON dokumentech by bylo dobré nastínit, jak taková data rozšířit, aby se z nich stala zároveň RDF data. Cílem je ale zachovat původní strukturu JSON souboru, resp. aby data byla validní vůči JSON schématu. Pro tyto účely je ideální formát JSON-LD. Jediné, co nám stačí, je v původních datech každému objektu přiřadit *@id*, *@type* a definovat *@context* (viz první kapitola Příklad serializovaných dat ve formátu JSON-LD). Při tvorbě ontologie jsme si popsali mapování entit a položek z datového standardu na konkrétní třídy a predikáty. V kontextu je tedy přesně takovéto mapování, viz příklad kódu 4.1. Na příkladu kódu 4.2 je již vidět výsledný JSON-LD soubor s jednou smlouvou a dvěma smluvními stranami (pro porovnání, původní JSON soubor viz kód 3.1). Jedná se tedy o RDF data, která popisuje námi definovaná ontologie a zároveň jde o data splňující datový standard, resp. jsou validní vůči JSON Schématu. Hlavním přínosem je to, že RDF data serializovaná v takto definovaném JSON-LD formátu budou použitelná v budoucích aplikacích, příp. registru pracujícím nad datovým standardem.

```

1 {
2   "@context": {
3
4     "cn": "http://tiny.cc/open-contracting#",
5     "com": "https://w3id.org/commerce#",
6     "dc": "http://purl.org/dc/terms/",
7     "dcmi": "http://dublincore.org/documents/dcmi-type-vocabulary/",
8     "foaf": "http://xmlns.com/foaf/0.1/",
9     "gr": "http://purl.org/goodrelations/v1#",
10    "owl": "http://www.w3.org/2002/07/owl#",
11    "ps": "https://w3id.org/payswarm#",
12    "schema": "http://schema.org/",
13    "rdf": "http://www.w3.org/1999/02/22-rdf-syntax-ns#",
14    "rdfs": "http://www.w3.org/2000/01/rdf-schema#",
15    "xsd": "http://www.w3.org/2001/XMLSchema#",
16
17    "id" : "dc:identifier",
18    "language" : "dc:language",
19
20    "Contract" : "cn:Contract",
21    "Attachment" : "cn:Attachment",
22    "Amendment" : "cn:Amendment",
23    "Order" : "cn:Order",
24    "Invoice" : "cn:Invoice",
25    "Publisher" : "foaf:Organization",
26    "Version" : "cn:Version",
27    "Party" : "gr:BusinessEntity",
28    "SuperiorInstitution" : "gr:BusinessEntity",
29    "Address" : "schema:PostalAddress",
30    "Implementation" : "cn:Implementation",
31    "Milestone" : "cn:Milestone",
32    "Transaction" : "com:Transaction",

```

```

33
34     "documents" : { "@id": "cn:documents", "@container": "@set" },
35     "orders" : { "@id": "cn:orders", "@container": "@set" },
36     "invoices" : { "@id": "cn:invoices", "@container": "@set" },
37
38     "uri" : { "@id": "cn:uri", "@type" : "@id" },
39     "document" : { "@id": "com:contentUrl", "@type" : "@id" },
40     "valid" : { "@id": "cn:valid", "@type" : "xsd:boolean" },
41     "anonymised" : { "@id": "cn:anonymised", "@type" : "xsd:boolean"
42     },
43     "responsiblePersons" : { "@id": "cn:responsiblePerson", "
44     @container": "@set" },
45     "attachments" : { "@id": "cn:attachment", "@container": "@set"
46     },
47     "amendments" : { "@id": "cn:amendment", "@container": "@set" },
48
49     "awardID" : "cn:awardID",
50     "awardProfileID" : "cn:awardProfileID",
51     "amount" : { "@id": "gr:hasCurrencyValue", "@type" : "xsd:float"
52     },
53     "amountNoVat" : { "@id": "gr:hasCurrencyValue", "@type" : "
54     xsd:float" },
55     "title" : "dc:title",
56     "type" : { "@id": "dcmi:type", "@container": "@set" },
57     "contractType" : "cn:contractType",
58     "subjectType" : "cn:subjectType",
59     "priceAnnual" : { "@id": "cn:priceAnnual", "@type" : "
60     xsd:boolean" },
61     "currency" : "gr:hasCurrency",
62     "dateSigned" : { "@id": "dc:created", "@type" : "xsd:date" },
63     "validFrom" : { "@id": "ps:validFrom", "@type" : "xsd:date" },
64     "validUntil" : { "@id": "ps:validUntil", "@type" : "xsd:date" },
65     "funding" : "cn:funding",
66     "currentValidContract" : { "@id": "cn:currentValidContract", "
67     @type" : "@id" },
68     "description" : "dc:description",
69     "competency" : { "@id": "cn:competency", "@container": "@set" },
70     "parties" : { "@id": "cn:party", "@container": "@set" },
71
72     "localID" : { "@id": "cn:localID", "@type" : "xsd:integer" },
73     "name" : "gr:legalName",
74     "payer" : { "@id": "cn:payer", "@type" : "xsd:boolean" },
75     "noID" : { "@id": "cn:noID", "@type" : "xsd:boolean" },
76     "country" : "schema:addressCountry",
77     "paysVAT" : { "@id": "gr:valueAddedTaxIncluded", "@type" : "
78     xsd:boolean" },
79     "superiorInstitution" : "cn:superiorInstitution",
80
81     "publisher" : "dc:publisher",
82     "authentication" : "cn:authentication",
83
84     "address" : "schema:address",
85     "streetAddress" : "schema:streetAddress",
86     "locality" : "schema:addressLocality",
87     "postalCode" : { "@id": "schema:postalCode", "@type" : "
88     xsd:integer" },
89     "nuts" : "cn:nuts",
90

```

```

82  "versions" : { "@id": "cn:version", "@container": "@set" },
83  "version" : { "@id": "cn:versionOrder", "@type": "xsd:integer"
},
84  "published" : { "@id": "dc:issued", "@type": "xsd:dateTime" },
85
86  "implementation" : "cn:implementation",
87  "milestones" : { "@id": "cn:milestone", "@type": "@id", "
@container": "@set" },
88  "transactions" : { "@id": "cn:transaction", "@type": "@id", "
@container": "@set" },
89  "dueDate" : { "@id": "cn:dueDate", "@type": "xsd:dateTime" },
90  "date" : { "@id": "dc:date", "@type": "xsd:dateTime" },
91  "publisherId" : "cn:publisherId",
92  "senderOrganization" : { "@id": "com:source", "@type": "@id" },
93  "receiverOrganization" : { "@id": "com:destination", "@type": "
@id" },
94
95  "contract" : { "@id": "cn:contract", "@type": "@id" },
96  "parrentDocument" : { "@id": "cn:parrentDocument", "@type": "
@id" }
97 }
98 }

```

Výpis kódu 4.1: JSON-LD Context

```

1  {
2    "@context": "http://tiny.cc/open-contracting-context",
3
4    "@id" : "http://rsmluv.cz/data/89f689cd-e784-4374-bb17-94144679
d46f",
5    "id": "89f689cd-e784-4374-bb17-94144679d46f",
6    "published": "2014-03-25T23:20:50+01:00",
7    "language": "cs",
8
9    "documents": [
10     {
11       "@id" : "http://rsmluv.cz/contract/12345/2",
12       "@type" : "Contract",
13       "uri" : "http://rsmluv.cz/contract/12345/2",
14       "document" : "http://rsmluv.cz/file/b15a3c45-5595-4a28-b156
-4578edeb2a98/Smlouva12345.docx",
15       "type" : "Smlouva",
16       "valid" : true,
17       "anonymised" : false,
18
19       "awardID": "486026",
20       "awardProfileID": "OI-010143",
21       "amount": 584520.00,
22       "title": "Brno, Vackova, Šafaříkova – rekonstrukce kanalizace
a vodovodu",
23       "contractType": "Kupní smlouva",
24       "subjectType": "Právní, finanční překladatelské, pojišťovnické
, poradenské a jiné služby",
25       "priceAnnual": false,
26       "currency": "CZK",
27       "dateSigned": "2011-11-16",
28       "validFrom": "2011-11-02",
29       "validUntil": "2012-06-30",
30       "competency" : [ "Soukromoprávní smlouva" ],

```

```

31     "currentValidContract": "http://zakazky.bрно.cz/?pg=detail&id
=18249&list=135",
32     "description": "Projektová dokumentace pro stavební povolení a
zadání stavby bude řešit rekonstrukci stávající kanalizační
stoky z profilu DN 500 na DN 800/1200 v délce 146 m, rekonstrukci
kanalizačních přípojek pod veřejným prostranstvím a přepojení vš
ech stávajících dešťových vpustí a také vybourání vozovek a chodn
íků nad rýhou a zásyp rýhy recyklátem. Součástí bude inženýrsko–
geologický průzkum, geodetické zaměření dotčené oblasti,
inventarizace zeleně, výkaz výměr, položkový rozpočet a výkon
autorského dozoru až do dokončení stavby. Dokumentace bude
projednána s orgány státní správy a s účastníky stavebního řízení
a jejich připomínky budou do dokumentace zapracovány.",
33
34     "responsiblePersons" : [ "Ing. Petr Vokřál ", "Mgr. Adriana
Krnáčová, MBA" ],
35
36     "publisher" : {
37         "@id" : "http://rsmluv.cz/contract/12345/2/publisher",
38         "@type" : "Publisher",
39         "id" : "6003508",
40         "name" : "Statutární město Brno",
41         "noID" : false,
42         "country" : "CZE",
43         "authentication" : "email"
44     },
45
46     "parties" : [
47         { "@id" : "http://rsmluv.cz/party/132456" },
48         { "@id" : "http://rsmluv.cz/party/987654" }
49     ],
50
51     "implementation" : {
52         "@id" : "http://rsmluv.cz/contract/12345/2/implementation",
53         "@type" : "Implementation",
54
55         "milestones" : [
56             {
57                 "@id" : "http://rsmluv.cz/contract/132456/2/milestone
/5830",
58                 "@type" : "Milestone",
59                 "title" : "Výpověď smlouvy",
60                 "dueDate" : "2012-06-20T23:20:50+01:00"
61             }
62         ],
63         "transactions" : [
64             {
65                 "@id" : "http://rsmluv.cz/contract/132456/2/transaction
/132456",
66                 "@type" : "Transaction",
67                 "publisherId" : "1269483",
68                 "date" : "2012-01-01T18:35:20+01:00",
69                 "amount" : 300000,
70                 "senderOrganization" : "http://rsmluv.cz/party/987654",
71                 "receiverOrganization" : "http://rsmluv.cz/party/987654"
72             }
73         ]

```



```

74         "@id" : "http://rsmluv.cz/contract/132456/2/transaction
/934584",
75         "@type" : "Transaction",
76         "publisherId" : "934584",
77         "date" : "2012-02-01T09:13:40+01:00",
78         "amount" : 284520,
79         "senderOrganization" : "http://rsmluv.cz/party/987654",
80         "receiverOrganization" : "http://rsmluv.cz/party/987654"
81     }
82 ]
83 },
84
85     "versions" : [
86     {
87         "version" : 1,
88         "@id" : "http://rsmluv.cz/contract/12345/1/version",
89         "@type" : "Version",
90         "uri" : "http://rsmluv.cz/contract/12345/1",
91         "published" : "2014-09-15T23:20:50+01:00"
92     },
93     {
94         "version" : 2,
95         "@id" : "http://rsmluv.cz/contract/12345/2/version",
96         "@type" : "Version",
97         "uri" : "http://rsmluv.cz/contract/12345/2",
98         "published" : "2015-03-15T14:35:28+01:00"
99     }
100 ]
101 }
102 ],
103
104     "parties" : [
105     {
106         "@id" : "http://rsmluv.cz/party/132456",
107         "@type" : "Party",
108         "localID" : 132456,
109         "id" : "44992785",
110         "name" : "Statutární město Brno",
111         "payer" : false,
112         "noID" : false,
113         "country" : "CZE",
114         "address" : {
115             "@id" : "http://rsmluv.cz/party/132456/address",
116             "@type" : "Address",
117             "streetAddress" : "Dominikánské náměstí 196/1",
118             "locality" : "Brno-město, Brno",
119             "postalCode" : 60200,
120             "nuts" : "CZ064"
121         },
122         "superiorInstitution" : {
123             "@id" : "http://rsmluv.cz/superiorInstitution/00064581",
124             "@type" : "SuperiorInstitution",
125             "id" : "00064581",
126             "localID" : 56486,
127             "name" : "Magistrát hlavního města Prahy",
128             "noID" : false,
129             "country" : "CZE"
130         }

```

```

131     },
132     {
133         "@id" : "http://rsmluv.cz/party/987654",
134         "@type" : "Party",
135         "localID" : 987654,
136         "id" : "46347011",
137         "name" : "Kovoprojekta Brno a.s.",
138         "payer" : true,
139         "noID" : false,
140         "country" : "CZE",
141         "address" : {
142             "@id" : "http://rsmluv.cz/party/987654/address",
143             "@type" : "Address",
144             "streetAddress" : "Šumavská 416/15",
145             "locality" : "Ponava, Brno",
146             "postalCode" : 60200,
147             "nuts" : "CZ064"
148         }
149     }
150 ]
151
152 }

```

Výpis kódu 4.2: JSON-LD Soubor s jednou smlouvou

## 5. Požadavky na platformu pro otevřené smlouvy

### 5.1 Funkční požadavky

### 5.2 Nefunkční požadavky

## 6. Návrh platformy pro otevřené smlouvy

### 6.1 Architektura

#### 6.1.1 Konverzní mechanismus

#### 6.1.2 Jednotné úložiště

#### 6.1.3 Webová aplikace

### 6.2 Linked Data v procesu otvírání smluv

## 7. Implementace platformy

### 7.1 Konverzní mechanismus

#### 7.1.1 Munis ESML

Struktura datového modelu

Omezení vůči standardu

#### 7.1.2 R2RML mapování

#### 7.1.3 Volba R2RML procesoru

Omezení R2RML procesoru

#### 7.1.4 Volba technologií a implementační platformy

#### 7.1.5 Napojení na datové úložiště

#### 7.1.6 SPARQL endpoint

#### 7.1.7 Zpracování RDF výstupu

Zpracování JSON-LD

### 7.2 Jednotné úložiště

#### 7.2.1 Nástroj Unified views

### 7.3 Webová aplikace

## 8. Zátěžové testy

8.1 Předpokládaná situace na úřadech

8.2 Výsledky

# Závěr

# Seznam použité literatury



# Seznam tabulek

# Seznam použitých zkratek

# Přílohy

# A Příloha

## B Příloha

## C Příloha

## D Příloha