

PŘEDNÁŠKA 8.

RDF a RDF schéma

Martina Husáková

3. 4. 2014

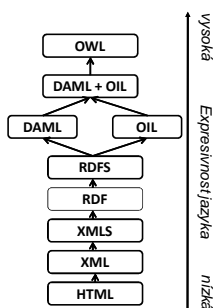
Obsah přednášky

- RDF
- Aplikace RDF
- RDF schéma

2

Ontologie RDF

- Resource Description Framework
- **Standard organizace W3C pro reprezentaci metadat o webových zdrojích v prostředí sémantického webu**
- Reprezentace vztahů mezi fakty, aby byly aplikace sémantického webu schopné data a informace snáze interpretovat



3

Cíle návrhu RDF ontologií

- Jednoduchý datový model
- Nabídka formální sémantiky a inference
- Použití rozšiřitelného slovníku založeného na URI
- Použití syntaxe založené na XML
- Komukoliv dovolit tvorbu tvrzení o jakémkoliv zdroji

[Zdroj: RDF: W3C Concepts and Abstract Syntax: <http://www.w3.org/TR/rdf11-concepts/>]

4

Využití RDF ontologií

- **Reprezentace metadat** o webových zdrojích
- Reprezentace informací o „věcech“, které mohou být identifikovatelné na webu
- Umožnit strojovou zpracovatelnost těchto informací o „věcech“ spíše než jejich prosté zobrazení člověku
- Možnost výměny RDF dat mezi různými aplikacemi bez ztráty významu dat (informací)

5

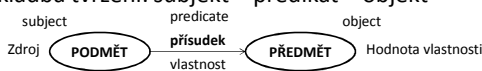
Standard RDF

- Ve formě standardu schválen r. 2004
- Klíčové dokumenty:
 - **RDF Concepts and Abstract Syntax**
 - RDF Semantics
 - **RDF Primer**
 - RDF/XML Syntax Specification
 - RDF Vocabulary Description Language 1.0
 - RDF Test Cases

6

RDF model

- RDF dekomponuje znalosti do malých fragmentů tzv. **trojic** (tripletů), které tvoří **tvrzení** vztahující se k „věcem“ (zdrojům) naší reality
- Skladba tvrzení: subjekt – predikát – objekt



- Primární vyjádření** RDF tvrzení: graf (uzel = podmět, předmět a hrana = přísudek)
- Sekundární vyjádření** RDF tvrzení: syntaxe

7

RDF model

- Zdroje (podměty) jsou popisovány vlastnostmi (přísudky) a jejich hodnotami (předměty)
- Podmět a předmět (podstatné jméno): názvy „věcí“ reálného světa
- Přísudek (sloveso): relace mezi těmito „věcmi“ (objektová vlastnost, vztah, asociace)

8

Analogie

- Databázová tabulka

Pořadí	Podmět	Přísudek	Předmět
Trojice 1.	Karel	<i>vlastní</i>	Notebook HP
Trojice 2.	Společnost HP	<i>je_výrobcem</i>	Notebook
Trojice 3.	Karel	<i>bydlí</i>	Hradec Králové

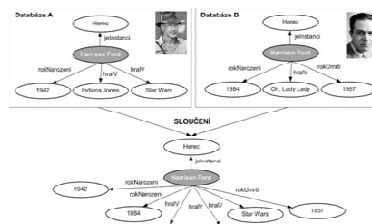
} tvrzení

- Součásti trojic tvořeny ve formě jednoduchých názvů a to „věcí“ konkrétních i abstraktních
- Názvy nemají žádnou vnitřní strukturu nebo stupeň priority
- Lze je chápat jako proměnné
- Názvy v RDF odkazují na „věci“ přítomné ve světě nebo je přímo reprezentují

9

Identifikace v RDF

- Není vhodné identifikovat „věci“ dle jména, zejména u většíchází dat při jejich slučování



[Zdroj: T. Segaran a kol. Programming the Semantic Web. O'Reilly Media, 2009.]

10

Identifikace v RDF

- RDF ontologie jsou založeny na myšlence identifikace „věcí“ s použitím identifikátorů
- V RDF se pro jednoznačnou identifikaci „věcí“ používá URI (Uniform Resource Identifier)
- Identifikátor může být přidělen subjektu, objektu nebo i predikátu

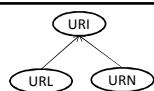
11

URI

- URI (Uniform Resource Identifier):** obecnější forma identifikátoru s možností adresace:
 - „věcí“ přístupných pomocí počítačové sítě (elektronické dokumenty, webové stránky, obrázky, videa, ...)
 - „věcí“ nepřístupných pomocí počítačové sítě (člověk, společnost, kniha, pes, ...)
 - abstraktních konceptů fyzicky neexistujících (štěstí, dynamika, stabilita, rovnováha, ...)
- Aplikace pro podmět, přísudek i předmět

12

URL a URN



- **URL (Uniform Resource Locator):** adresace zdrojů získatelných z webu přímo pomocí řetězce, který zdroj identifikuje (např. <http://www.w3.org/TR/rdf-primer/>)
- **URN (Uniform Resource Name):** Jednoznačně určuje název zdroje, dle kterého by ho měl klient rozpoznat, ale neříká, jak jej získat (obdoba ISBN k identifikaci publikací)

13

Reprezentace tvrzení: Příklad 1

Tvrzení 1.

Autorem webové stránky <http://www.example.org/index.html> je John Smith.

Tvrzení 2.

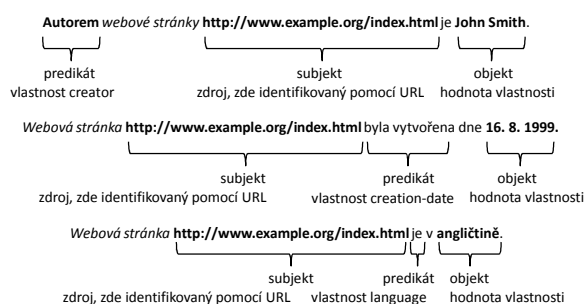
Webová stránka <http://www.example.org/index.html> byla vytvořena dne 24. 12. 2011.

Tvrzení 3.

Webová stránka <http://www.example.org/index.html> je v angličtině.

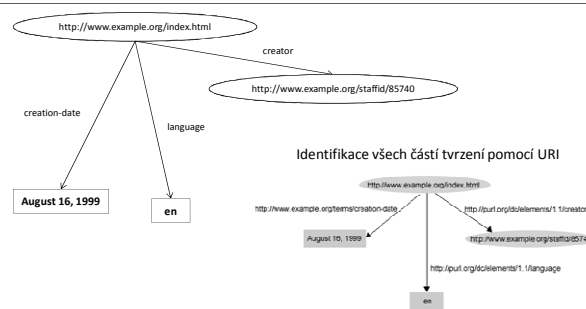
14

Reprezentace tvrzení: příklad 1



15

Příklad 1 ve formě grafu



[Zdroj: RDF Primer – Recommendation, 2004. URL: <http://www.w3.org/TR/rdf-primer/>]

16

Reprezentace tvrzení

- Subjekt a predikát: ve formě URI reference
- Objekt: ve formě URI reference nebo konstantní hodnoty (tzv. literálu)
- Literál:
 - Typový literál: s udáním datového typu
"27"^^xsd:integer, "1999-08-08"^^xsd:date
 - Netypový literál (plain): bez udání datového typu
27, en, surname
 - Neslouží k reprezentaci subjektů a predikátů!

17

Alternativní zápis tvrzení: příklad 1

- Ne vždy je RDF graf vhodný pro reprezentaci tvrzení (strojové zpracování)



N-Triples notation (Triples notation)

<<http://www.example.org/index.html> <<http://purl.org/dc/elements/1.1/creator>> <<http://www.example.org/staffid/85740>>

<<http://www.example.org/index.html> <<http://www.example.org/terms/creation-date>> "August 16, 1999"

<<http://www.example.org/index.html> <<http://purl.org/dc/elements/1.1/language>> "en"

[Zdroj: RDF Primer – Recommendation, 2004. URL: <http://www.w3.org/TR/rdf-primer/>]

18

Prefixy aneb zjednodušíme zápis

- Přiřazování prefixů jmenným prostorům je běžné
- Prefix zastupuje jmenný prostor
- Není třeba specifikovat celou webovou adresu

prefix rdf:, namespace URI: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
 prefix rdfs:, namespace URI: <http://www.w3.org/2000/01/rdf-schema#>
 prefix dc:, namespace URI: <http://purl.org/dc/elements/1.1/>
 prefix owl:, namespace URI: <http://www.w3.org/2002/07/owl#>
 prefix ex:, namespace URI: <http://www.example.org/> (or <http://www.example.com/>)
 prefix xsd:, namespace URI: <http://www.w3.org/2001/XMLSchema#>

19

Zkrácení zápisu: příklad 1

Navržené prefixy

prefix ex:, namespace URI: <http://www.example.org>
 prefix exstems:, namespace URI: <http://www.example.org/terms/>
 prefix exstaff:, namespace URI: <http://www.example.org/staff/>
 prefix dc:, namespace URI: <http://purl.org/dc/elements/1.1/>

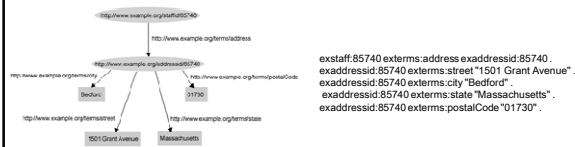
ex:index.html dc:creator exstaff:85740
 ex:index.html exstems:creation-date "August 16, 1999"
 ex:index.html dc:language "en"

[Zdroj: RDF Primer – Recommendation, 2004. URL: <http://www.w3.org/TR/rdf-primer/>]

20

Rozšíření příkladu 1

- Adresa Johna Smithe
- Možnost A
- Možnost B: adresa strukturovaněji, ale s novým uzlem (exaddressid:85740)

[Zdroj: RDF Primer – Recommendation, 2004. URL: <http://www.w3.org/TR/rdf-primer/>]

21

Typové a netypové literály

- Tvrzení: Johnu Smithovi je 27 let.

Reprezentace netypovým literálem

<http://www.example.org/staff/85740> <http://www.example.org/terms/age> 27

Reprezentace typovým literálem

<http://www.example.org/staff/85740>
 <http://www.example.org/terms/age>
 "27"^^<http://www.w3.org/2001/XMLSchema#integer>

Reprezentace typovým literálem (QName)

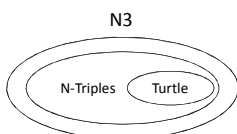
exstaff:85740 exstems:age "27"^^xsd:integer

[Zdroj: RDF Primer – Recommendation, 2004. URL: <http://www.w3.org/TR/rdf-primer/>]

22

Syntaxe pro RDF

- RDF graf je vhodnou reprezentací informací v příp. jednoduchých tvrzení
- Komplexní tvrzení, čítající více než 100 uzlů, jsou problematická ke zpracování i člověkem
- Syntaxe pro RDF:
 - Notace 3 (N3)
 - N-Triple
 - Turtle
 - RDF/XML



[Zdroj: Hitzler, P. a kol. Foundation of semantic web technologies. Chapman and Hall/CRC, 1 edition, 2009]

23

RDF aplikace: FOAF

- FOAF (angl. Friend Of A Friend)
 - slovník pro popis osob a vztahů mezi nimi reprezentovaný pomocí RDF
 - (... zaměstnání osob; názvy institucí; projekty, na kterých osoba pracuje; dokumenty vytvořené danou osobou; fotografie osoby, ...)

```
<#me>
a foaf:Person ;
foaf:name "Jiří Grygar" ;
foaf: mbox mailto:grygar@fzu.cz ;
foaf:knows [
  a foaf:Person ;
  foaf:name "Josip Kleczek" ;
  foaf: mbox mailto:grygar@fzu.cz
]
```

„Koho zná Jiří Grygar?“
Zápis v Turtle syntaxi

[Zdroj: RDF Primer – Recommendation, 2004. URL: <http://www.w3.org/TR/rdf-primer/>]

24

Vlastnosti RDF (1)

- Lze tvořit jen jednoduchá tvrzení tvaru
podmět-přísudek-předmět
– Krkonoše *má* Nejvyšší Horu Sněžka
- Nelze vyjádřit co podmět, přísudek a předmět
znamenají, ani jaké jsou mezi nimi vztahy
– Krkonoše jsou pohoří, Sněžka je hora, každá hora
má výšku, ...
- Pro tvorbu komplexních ontologií se nehodí,
vhodnější je OWL

25

Vlastnosti RDF (2)

- Explicitně nemá zabudován koncept
reprezentace tříd, implicitně ano ve formě
rdf:type a *rdf:property* => v RDF lze
reprezentovat, že „něco“ je třídou nebo
vlastností
- S RDF ale nelze reprezentovat, že třída A je
podtřídou třídy B => hierarchii tříd a vztahy
mezi třídami

26

RDF schéma

- RDF Schema, The RDF Vocabulary Description language
- Nadstavba jazyka RDF pro tvorbu ontologií, která vznikla rozšířením
syntaxe a sémantiky RDF
- Určeno pro tvorbu struktury RDF tvrzení, resp. pro tvorbu tříd
zdrojů a tříd vlastností pro reprezentaci vztahů mezi třídami
- Zdroje tak mohou být kategorizovány a vnímány jako instance tříd
- Třídy zdrojů jsou seskupovány do hierarchie
- Třídy jsou vnímány jako rozšíření popisu zdrojů, více je konkretizují
- Hierarchie RDF tříd je chápána jako RDF slovník obsahující množinu
předdefinovaných RDF tvrzení se specifickým významem
- Jedná se také o RDF graf, jako v případě RDF
- Jmenný prostor má následující tvar:
<http://www.w3.org/2000/01/rdf-schema#> (rdfs prefix)

[Zdroj: V. Kvasnička, a kol. Umele inteligence a kognitivní veda II. STU, 2010. ISBN 978-80-227-3284-0.]

27

RDF Schéma: Co nabízí navíc?

- Specifikace definičního oboru a oboru hodnot
(*rdfs:domain*, *rdfs:range*)
- Specifikace datových typů (*rdfs:Datatype*)
- Reprezentace tříd a jejich hierarchií (*rdfs:Class*,
rdfs:subClassOf)
- Reprezentace hierarchie vlastností
(*rdfs:subPropertyOf*)
- Tvorba komentářů (*rdfs:comment*)
- ...

[Zdroj: V. Kvasnička, a kol. Umele inteligence a kognitivní veda II. STU, 2010. ISBN 978-80-227-3284-0.]

28

Příklad: tvorba hierarchie tříd

- Třída je jakýkoliv zdroj, který má specifikovanou
vlastnost (*rdf:type*), jejíž hodnotou je právě
rdfs:Class
- Reprezentace třídy Auto
– Ex:Auto *rdf:type* *rdfs:Class*
- Reprezentace třídy DopravníProstředek
– Ex:DopravníProstředek *rdf:type* *rdfs:Class*
- Reprezentace třídy Auto jako podtřídy třídy
DopravníProstředek
– Ex:Auto *rdf:subClassOf* *rdfs:DopravníProstředek*

Poznámka: Ex: URIref pro <http://www.example.org/schemas/vehicles>

29

Příklad RDFS grafu

[Zdroj: RDF Primer: W3C Recommendation, 2004. URL: <http://www.w3.org/TR/rdf-primer/>]

30

Příklad: tvorba vlastností



- Relace mezi třídami pomocí třídy **rdf:Property**
- Reprezentace vlastnosti weightInKg
`ex:terms:weightInKg rdf:type rdf:Property`
- Definiční obor objektové vlastnosti weightInKg
 - Ex: `weightInKg rdfs:domain ex:Motor`
- Obor hodnot objektové vlastnosti weightInKg
 - Ex: `weightInKg rdfs:range ex:Weight`

31

Silné stránky RDF schématu



- Tvorba jednoduchých ontologií, resp. vztahů mezi třídami
- Nízká výpočetní složitost
- S RDF schématem lze vyjádřit:
 - explicitně třídu,
 - datové typy,
 - def. obor, obor hodnot,
 - podtřídu,
 - „pod-vlastnost“, a další ...

32

Slabé stránky RDFS



- Nelze vyjádřit (vymezit):
 - ekvivalenci tříd
 - inverzní vlastnosti
 - funkcionalitu vlastnosti nebo její tranzitivitu
 - omezení kardinality vlastností
 - negaci tvrzení
 - disjunktnost tříd, ale pouze vztah je podtřídou
 - třídu použitím logických operátorů AND (průnik), OR (sjednocení), NOT (negace) na třídy
- Nevhodné pro tvorbu komplexních ontologií
 OWL je vhodnější volbou

33

RDF schéma vs. OWL



- OWL si bere z RDF schopnost tvořit strukturu tříd a vlastností, přidává sémantiku

Možnosti RDFS i OWL	Možnosti OWL, nikoliv RDFS
Deklarace tříd a podtříd	Disjunktnost tříd
Reprezentace jedinců tříd	DifferentFrom, SameAs
Reprezentace definičního oboru a oboru hodnot (třídy)	Inverzní vlastnost, tranzitivita vlastnosti, symetrie vlastnosti, funkcionalita a kardinalita vlastnosti
Reprezentace definičního oboru (třída) a oboru hodnot (dato-typová hodnota)	Množinové operace
Jedinec má vlastnost s určitou dato-typovou hodnotou	

[Zdroj: I. Horrocks, a kol. From SHIQ and RDF to OWL: The Making of a Web Ontology Language. 2003.]

34

Samostudium

Příští přednáška dne 10. 4. 2014 na téma

Rozšíření a alternativy k RDF

35