

4IZ440 Propojená data na webu

RDF serializace Turtle

Vyučující: prof. Ing. Vojtěch Svátek, Dr.

Zimní semestr 2018

<http://nb.vse.cz/~svatek/rzzw.html>

Shrnutí obsahu a související zdroje

- Prezentace vychází ze starší [prezentace](#) (2012)
J. Mynarze
 - Některé údaje jsou aktualizované
 - Jsou vynechány některé části, které vyučující v současnosti považuje buď za méně podstatné nebo kolidující s jinými materiály: kolekce, část „RDF slabikáře“
- Hlavní rozšiřující zdroj je specifikace Turtle:
<https://www.w3.org/TR/turtle/>
(doporučení W3C z února 2014)
- Prezentace bude v případě potřeby aktualizována, např. o komentáře k této specifikaci

Turtle

- Jedna ze serializací datového modelu RDF
- Dalšími jsou zejména:
 - **RDF/XML**: první standardizovaná, rozšířená; špatně čitelná, vrstva XML je často nadbytečnou zátěží při zpracování
 - **N-Triples**: jednoduchá, dobře se zpracovává i komprimuje pro přenos; kvůli absenci zkracování špatně čitelná
 - **RDFa**: lze embedovat do HTML stránek a XML dokumentů
 - **JSON-LD**: rovněž umožňuje embedování, obliba u webových vývojářů
 - **Notation 3 (N3)**: vyvinul TBL, nadstavba Turtle umožňující mj. pravidlové odvozování; malá míra využívání
 - **TriG** a **N-Quads**: rozšíření Turtle resp. N-Triples pro práci s pojmenovanými grafy
- Konverze serializací viz např. <http://any23.org/>

Charakteristiky Turtle

- Turtle = Terse RDF Triple Language
- Znakové kódování UTF-8
- MIME type: `text/turtle`
- Přípona souborů: `.ttl`
- Hlavní užití
 - **Manuální** vytváření RDF
 - Grafové vzory v dotazovacím jazyce **SPARQL**

N-Triples vs. Turtle vs. RDF/XML

- **N-Triples (odřádkování je až za tečkou!):**

<http://keg.vse.cz/resource/person/jindrichmynarz>

<http://xmlns.com/foaf/0.1/knows> <http://keg.vse.cz/resource/person/vojtech-svatek> .

<http://keg.vse.cz/resource/person/jindrichmynarz>

<http://xmlns.com/foaf/0.1/name> "Jindřich Mynarz" .

- **Turtle (využití prefixů a UTF-8; zatím bez dalšího zkracování):**

@prefix foaf: <http://xmlns.com/foaf/0.1/> .

@prefix person: <http://keg.vse.cz/resource/person/> .

person:jindrich-mynarz foaf:knows person:vojtech-svatek .

person:jindrich-mynarz foaf:name "Jindřich Mynarz" .

- **RDF/XML:**

<?xml version="1.0" encoding="utf-8"?>

<rdf:RDF xmlns:foaf="http://xmlns.com/foaf/0.1/" xmlns:person="http://keg.vse.cz/resource/person/"

xmlns:rdf="http://www.w3.org/1999/02/22-rdfsyntax-ns#">

<rdf:Description rdf:about="http://keg.vse.cz/resource/person/jindrich-mynarz">

<foaf:knows rdf:resource="http://keg.vse.cz/resource/person/vojtech-svatek"/>

<foaf:name>Jindřich Mynarz</foaf:name>

</rdf:Description> </rdf:RDF>

Prvky Turtle (a obecně RDF trojic)

- **IRI**: na jakékoli pozici
- **Literály**: jen na pozici objektu
- **Lokální identifikátory** („blank nodes“):
na pozici subjektu nebo objektu

IRI

- Internationalized Resource Identifier
 - Vzniklo z URI (Uniform Resource Identifier) rozšířením o znaky jiných abeced
- Dva typy zápisu
 - Absolutní IRI: `<http://example.com/uri>`
 - Relativní IRI: `<uri>`
 - Odkazuje na IRI **výchozího** jmenného prostoru:
`@base <http://example.com/>` .

Literály

- Jednoduché
 - string: `"Literál"`
 - integer: `12345`
 - double: `4.2E9`
 - decimal: `-5.0`
 - boolean: `true` nebo `false`
- Anotované
 - Jazykovým tagem: `"Literál"@cs`
 - Datový typem: `"Literál"^^xsd:string`

Lokální identifikátory

- Též „blank nodes“ (zakreslované jako „prázdné uzly“ v RDF grafu), „anonymní uzly“
- Možné zápisy (zjednodušeně):
 - `_:identifikátorBlankNode`
 - `[]` (hrnaté závorky) - varianta používaná zejména při zkracování, do závorek lze vepsat zbytek trojice (predikát a objekt) za blank nodem, který je subjektem, viz dále
- Mají platnost jen v rámci **dané množiny RDF trojic** (dokumentu obsahujícího serializovaná data; případně grafu v úložišti RDF npod.), nelze na ně odkazovat zvenku
- Využívání v **datech** spíše nedoporučováno, hlavní smysl mají ve SPARQL **dotazech**

Další prvky Turtle

- Oddělovače
 - Subjekt, predikát a objekt se oddělují bílými znaky
 - Trojice (jakoby „věty“ dokumentu) se oddělují tečkou
- Komentáře
 - # komentář

Zkracování: jmenné prostory

- Deklarace **prefixů** jmenných prostorů pomocí `@prefix`
 - Např. přiřazení prefixu "rdf" k IRI:
`@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .`
 - IRI může být i relativní:
`@base <http://keg.vse.cz/resource/> .`
`@prefix person: <person/> .`
`@prefix project: <project/> .`
- Na prefix následně odkazují **prefixovaná jména**, („kompaktní URI, CURIE“) např.:
 - `rdf:type`
 - `rdf` = IRI prefix, `type` = lokální jméno (IRI fragment)
 - `person:jindrich-mynarz`, `project:openbudgets`
- Obvykle užívané prefixy viz <http://prefix.cc>

Zkracování: společný subjekt

- Páry predikát-objekt, které náležejí ke **stejnému subjektu**, oddělujeme středníkem (";")

Místo

@prefix foaf: <http://xmlns.com/foaf/0.1/> .

@prefix person: <http://keg.vse.cz/resource/person/> .

person:jindrich-mynarz foaf:knows person:vojtech-svatek .

person:jindrich-mynarz foaf:name "Jindřich Mynarz" .

zapišeme

@prefix foaf: <http://xmlns.com/foaf/0.1/> .

@prefix person: <http://keg.vse.cz/resource/person/> .

person:jindrich-mynarz foaf:knows person:vojtech-svatek ;

foaf:name "Jindřich Mynarz" .



Zkracování: seznam objektů

- Objekty, které jsou všechny hodnotou **stejného predikátu** (pro daný subjekt), oddělujeme čárkou (,,")

Místo

@prefix foaf: <http://xmlns.com/foaf/0.1/> .

@prefix person: <http://keg.vse.cz/resource/person/> .

person:jindrich-mynarz foaf:knows person:vojtech-svatek .

person:jindrich-mynarz foaf:knows person:ondrej-zamazal .

zapišeme

@prefix foaf: <http://xmlns.com/foaf/0.1/> .

@prefix person: <http://keg.vse.cz/resource/person/> .

person:jindrich-mynarz foaf:knows person:vojtech-svatek ,
person:ondrej-zamazal .



Zkracování: blank nodes

Místo

@prefix foaf: <http://xmlns.com/foaf/0.1/> .

@prefix person: <http://keg.vse.cz/resource/person/> .

person:jindrich-mynarz foaf:knows _:tomas-knap .
_:tomas-knap foaf:name "Tomáš Knap".

zapišeme s využitím hranatých závorek

@prefix foaf: <http://xmlns.com/foaf/0.1/> .

@prefix person: <http://keg.vse.cz/resource/person/> .

person:jindrich-mynarz foaf:knows [foaf:name "Tomáš Knap"] .



Zkracování: `rdf:type`

- Predikát `rdf:type` (vztah zdroje a jeho třídy) se zkracuje na `a` (ve smyslu "is a")

Místo

@prefix foaf: <http://xmlns.com/foaf/0.1/> .

@prefix person: <http://keg.vse.cz/resource/person/> .

@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .

person:jindrich-mynarz rdf:type foaf:Person .

zapišeme

@prefix foaf: <http://xmlns.com/foaf/0.1/> .

@prefix person: <http://keg.vse.cz/resource/person/> .

@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .

person:jindrich-mynarz a foaf:Person .



Zkracování: cvičení 1

<<http://www.w3.org/People/Berners-Lee/card#i>>

<<http://www.w3.org/1999/02/22-rdf-syntax-ns#type>>

<<http://xmlns.com/foaf/0.1/Person>>

.

Zkracování: cvičení 2

<<http://www.w3.org/People/Berners-Lee/card#i>>

<<http://www.w3.org/1999/02/22-rdf-syntax-ns#type>>

<<http://xmlns.com/foaf/0.1/Person>>

.

<<http://www.w3.org/People/Berners-Lee/card#i>>

<<http://xmlns.com/foaf/0.1/knows>>

<<http://richard.cyganiak.de/foaf.rdf#cygri>>

.

<<http://www.w3.org/People/Berners-Lee/card#i>>

<<http://xmlns.com/foaf/0.1/knows>>

<<http://mhausenblas.info/#i>>

.

Zkracování: cvičení 3

<http://example.com/day1>

<http://example.com/property/temperature>

"-0.3"^^<http://www.w3.org/2001/XMLSchema#decimal>

.