

RDF

Лекција 5

Стеван Гостојић

Факултет техничких наука, Нови Сад

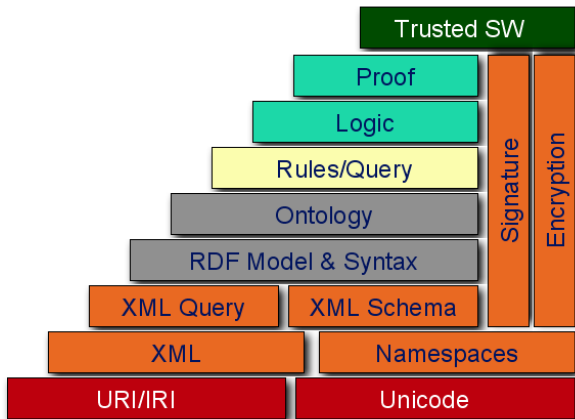
21. март 2024.

Преглед садржаја

- 1 Увод
- 2 Resource Description Framework
- 3 SPARQL
- 4 Закључак

- W3C (World Wide Web Consortium) је непрофитна организација која се бави стандардизацијом World Wide Web-а
- W3C препоруке (стандарди) су јавно и бесплатно доступни на <https://www.w3.org/>

Semantic Web стек





*The XPS and Semantic Web (Worlds) Technologies, Interoperability and Integration: A survey of the State of the Art in the Semantic Web (Multi-modal Adaptive Systems and Applications), Springer 2009.

Ресурси

- Ресурс је било шта што има идентитет (RFC 2396)
 - (електронски) документи
 - сервиси
 - колекција ресурса

Репрезентације ресурса

- Репрезентација ресурса је информација која рефлектује прошло, тренутно или жељено стање ресурса, у формату који може да се лако комуницира преко протокола, и која се састоји од скупа репрезентационих метаподатака и потенцијално неограниченим током репрезентационих података (RFC 7231)

Репрезентације ресурса

- Сваки ресурс може да има више репрезентација:
 - (X)HTML
 - XML
 - RDF
 - JSON
 - итд.

URI

- URI (Uniform Resource Identifier) је низ знакова који идентификује апстрактне или физичке ресурсе (RFC 2396)
- Јавља се у два облика: Uniform Resource Locator (URL) и Uniform Resource Name (URN)

URL

- Uniform Resource Locator (URL) је подскуп URI који идентификује ресурсе преко примарног механизма приступа (нпр. преко локације на мрежи)
 - <http://www.ftn.uns.ac.rs/>
 - <ftp://ftp.is.co.za/rfc/rfc1808.txt>
 - <mailto:mduerst@ifi.unizh.ch>

URN

- Uniform Resource Name (URN) је подскуп URI који је глобално јединствен и перзистентан чак иако ресурс постане недоступан или престане да постоји
 - urn:isbn:0-395-36341-1
 - urn:uuid:f81d4fae-7dec-11d0-a765-00a0c91e6bf6
 - urn:lex:eu:commission:directive:2010-03-09;2010-19-EU

Преглед садржаја

- 1 Увод
- 2 Resource Description Framework
- 3 SPARQL
- 4 Закључак

Повезани подаци

- Повезани подаци (енг. linked data) су скуп принципа за објављивање, проналажење и прегледање података у RDF формату (који могу да буду дистрибуирани на више сервера)

Принципи повезаних података (Tim Berners-Lee 2007)

- Користити URI за имена ствари
- Користити HTTP URI да би ствари могле да се пронађу
- Када неко затражи ресурс, понудити му репрезентацију ресурса у RDF формату
- У одговор укључити RDF исказе који повезују тражени ресурс са другим ресурсима (да би могле да се пронађу сродне ствари)

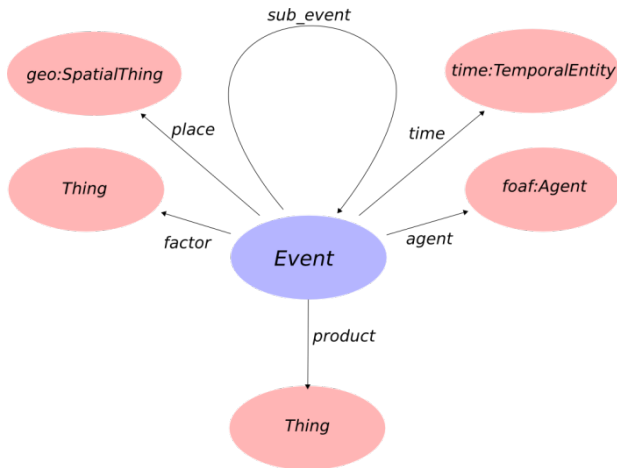
RDF

- Resource Description Framework (RDF) је стандардни модел за размену података на вебу
- Проширује људски читљиву мрежу докумената са машински читљивом мрежом података

RDF

- Подаци чине усмерен означен граф
- Чворови графа су ствари и типизирани константе, а гране графа су релације између ствари или релације између ствари и типизираних константи
- Ова структура може се представити као скуп исказа облика "<subjekat> <predikat> <objekat> ." (тј. скуп уређених тројки)

RDF граф



RDF граф

```
1 @prefix event: <http://purl.org/NET/c4dm/event.owl#> .
2 @prefix foaf: <http://xmlns.com/foaf/0.1/> .
3
4 Event
5   event:factor Thing ;
6   event:agent foaf:Agent ;
7   event:place <http://sws.geonames.org/2643744/> ;
8   event:time time:TemporalEntity .
9
```

RDF

- RDF унифицира начин на који се описују ресурси (тј. начин на који се представљају информације)

HTML

```
1 <!DOCTYPE html>
2 <html xmlns="http://www.w3.org/1999/xhtml">
3   <body>
4     <p>William Shakespeare is author of Hamlet.</p>
5   </body>
6 </html>
7
```

HTML

```
1 <!DOCTYPE html>
2 <html xmlns="http://www.w3.org/1999/xhtml">
3   <body>
4     <h1>Hamlet</h1>
5     <p>Author: William Shakespeare</p>
6   </body>
7 </html>
8
```

XML

```
1 <author>
2   <book>Hamlet</book>
3   <name>William Shakespeare</name>
4 </author>
5
```

XML

```
1 <book>
2   <name>Hamlet</name>
3   <author>William Shakespeare</author>
4 </book>
5
```

XML

```
1 <book>
2   <name>Hamlet</name>
3   <author>
4     <name>William Shakespeare</name>
5   </author>
6 </book>
7
```


RDF

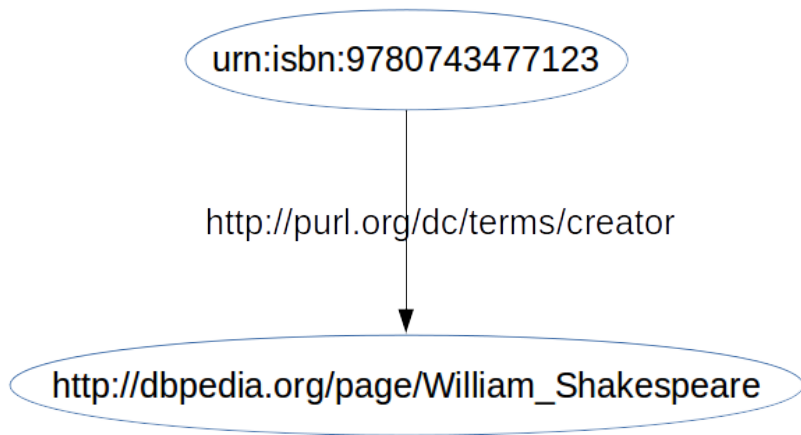


Figure 1: RDF iskaz.

RDF

```
1 <urn:isbn:9780743477123>  
2   <http://purl.org/dc/elements/1.1/creator>  
3     <http://dbpedia.org/page/William_Shakespeare> .  
4
```

RDF исказ

- Сваки RDF исказ састоји се од:
 - субјекта (ресурса)
 - предиката (својства)
 - објекта (ресурса или литерала)

RDF исказ

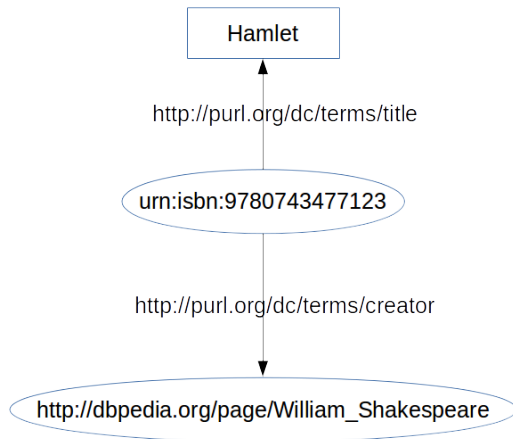


Figure 2: RDF исказ.

Ресурси

- Ресурси су било шта што има идентитет
- То су ствари које се описују са RDF
- Ресурси се идентификују са URI (URL или URN)

Својства

- Својства су специфични аспекти (карактеристике, атрибути итд.) који се користе за опис ресурса
- Вредност својстава може да буде ресурс или литерал (типизирана константа)
- Својства се идентификују са URI (URL или URN)

Литерали

- Литерали су (типизирани) константе
- XML Schema типови података
- ISO кодови језика

RDF граф

- Један или више RDF исказа чини један RDF граф
- Граф је уређени пар (V, E)
- V је скуп чворова (ресурса и литерала)
- E је скуп грана или релација над скупом чворова (својстава)

RDF граф

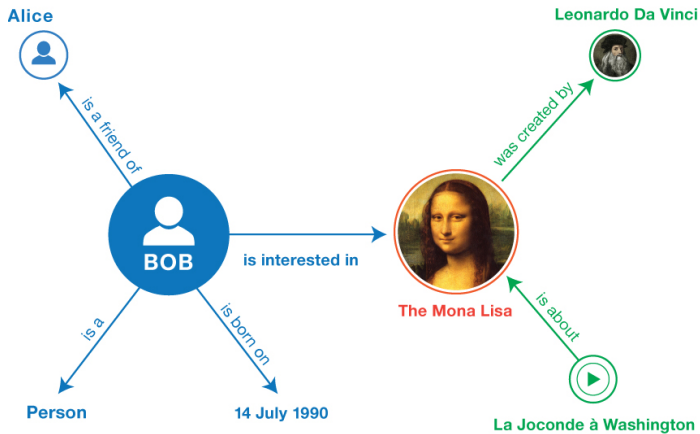


Figure 3: RDF граф.

RDF граф

```

1  BASE    <http://example.org/>
2  PREFIX  foaf: <http://xmlns.com/foaf/0.1/>
3  PREFIX  xsd: <http://www.w3.org/2001/XMLSchema#>
4  PREFIX  schema: <http://schema.org/>
5  PREFIX  dcterms: <http://purl.org/dc/terms/>
6  PREFIX  wd: <http://www.wikidata.org/entity/>
7
8  <#bob>
9    a foaf:Person ;
10    foaf:knows <#alice> ;
11    schema:birthDate "1990-07-14"^^xsd:date ;
12    foaf:topic_interest wd:Q12418 .
13
14  wd:Q12418
15    dcterms:title "Mona Lisa" ;
16    dcterms:creator <http://dbpedia.org/resource/Leonardo_da_Vinci> .
17
18  <http://data.europeana.eu/item/04802/243
19    FA8618938F4117025F17A8B813C5F9AA4D619>
20    dcterms:subject wd:Q12418 .

```

RDF формати

- Turtle
- RDF/XML
- RDFa
- итд.

Turtle

- Turtle (Terse RDF Triple Language) је конкретна синтакса за RDF
- Другим речима, то је један тип текстуалне серијализације RDF графа
- Акценат је на концизности и лакој читљивости

Turtle

```
1 @base <http://example.org/> .
2 @prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
3 @prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
4 @prefix foaf: <http://xmlns.com/foaf/0.1/> .
5 @prefix rel: <http://www.perceive.net/schemas/relationship/> .
6
7 <#green-goblin>
8     rel:enemyOf <#spiderman> ;
9     a foaf:Person ;      # in the context of the Marvel universe
10    foaf:name "Green Goblin" .
11
12 <#spiderman>
13     rel:enemyOf <#green-goblin> ;
14     a foaf:Person ;
15     foaf:name "Spiderman" .
16
```

Turtle

```
1 <http://example.org/#spiderman>  
2   <http://www.perceive.net/schemas/relationship/enemyOf>  
3     <http://example.org/#green-goblin> .  
4
```

Turtle

```
1 <http://example.org/#spiderman>
2   <http://www.perceive.net/schemas/relationship/enemyOf>
3     <http://example.org/#green-goblin> ;
4   <http://xmlns.com/foaf/0.1/name> "Spiderman" .
5
```

Turtle

```
1 <http://example.org/#spiderman>
2   <http://www.perceive.net/schemas/relationship/enemyOf>
3     <http://example.org/#green-goblin> ;
4   <http://xmlns.com/foaf/0.1/name>
5     "Spiderman"^^<http://www.w3.org/2001/XMLSchema#string>.
6
```


Turtle

```
1 <http://example.org/#spiderman>
2   <http://xmlns.com/foaf/0.1/name> "Spiderman"@en, "Spajdermen"@sr
3   .
```

Turtle

```
1 @base <http://example.org/> .  
2 @prefix rel: <http://www.perceive.net/schemas/relationship/> .  
3  
4 <#green-goblin> rel:enemyOf <#spiderman> .  
5
```

RDF/XML

- RDF/XML је конкретна синтакса за RDF
- Заснована је на XML-у (тј. то је један дијалекат XML-а)
- Акценат је на максималној интероперабилности

RDFa

- RDFa (RDF in attributes) је конкретна синтакса за RDF
- Заснована је на уграђивању RDF исказа у XML документе коришћењем стандардних атрибута

Преглед садржаја

- 1 Увод
- 2 Resource Description Framework
- 3 SPARQL**
- 4 Закључак

SPARQL

- SPARQL је упитни језик за постављање упита над подацима складиштеним у RDF формату
- SPARQL је комуникациони протокол за постављање упита над удаљеним RDF складиштима података (преко HTTP протокола)

RDF тројке

```
1 PREFIX dc: <http://purl.org/dc/elements/1.1/>
2
3 <urn:isbn:9780743477123> dc:title "Hamlet" .
4 <urn:isbn:9780743477116> dc:title "Romeo and Juliet" .
5 <urn:isbn:9780743477550> dc:title "Othello" .
6 <urn:isbn:9780743477109> dc:title "Macbeth" .
7 <urn:isbn:9780743482837> dc:title "The Tempest" .
8
```

SPARQL упит

```
1 PREFIX dc: <http://purl.org/dc/elements/1.1/>
2
3 SELECT ?title
4 WHERE {
5   <urn:isbn:9780743477123> dc:title ?title
6 }
7
```


SPARQL резултати упита

```
1  <?xml version="1.0"?>
2  <sparql xmlns="http://www.w3.org/2005/sparql-results#">
3    <head>
4      <variable name="title" />
5    </head>
6
7    <results>
8      <result>
9        <binding name="title">Hamlet</binding>
10      </result>
11    </results>
12  </sparql>
13
```

Извори података

- SPARQL упити се извршавају над RDF графовима
 - подацима изворно складиштеним у RDF формату
 - подацима складиштеним у другим форматима (XML, HTML, RDB) које middleware трансформише у RDF формат

Врсте SPARQL упита

Врста	Опис
SELECT	Враћа резултате у облику табеле (релације)
CONSTRUCT	Враћа резултате у облику (RDF) графа
ASK	Враћа boolean вредност (у зависности да ли еквивалентан SELECT упит враћа резултат)
DESCRIBE	Враћа (RDF) граф који описује ресурс

Table 1: Врсте SPARQL упита

Структура SPARQL упита

```
1  # prefix declarations
2  PREFIX foo: <http://example.com/resources/>
3  ...
4  # datasource definition
5  FROM <http://example.com/resources/>
6  ...
7  # result clause
8  SELECT ...
9  # graph pattern
10 WHERE {
11   ...
12 }
13 # query modifiers
14 ORDER BY ...
15
```

Структура SPARQL упита

- Декларација префикса (скраћивање URI)
- Дефиниција извора података (RDF графова над којима се извршава упит)
- Клаузула резултата (спецификација података које упит враћа)
- Образац графа (спецификација рестрикција над подацима)
- Модификатори упита (сортирање, груписање, исецање итд.)

Образац графа

- SPARQL променљиве почињу са "?" и могу да се поклапају са било којим чвором или граном у RDF графу
- Образац тројке је тројка у којој је бар један елемент (субјекат, предикат или објекат) замењен са променљивом
- Образац графа је скуп образаца тројки (који могу да буду филтрирани логичким изразима) и модификовани (OPTIONAL и UNION модификаторима)

SELECT

```
1 PREFIX foaf: <http://xmlns.com/foaf/0.1/>
2
3 SELECT ?name
4 WHERE {
5     ?person foaf:name ?name .
6 }
7
```

SELECT DISTINCT

```
1  PREFIX foaf: <http://xmlns.com/foaf/0.1/>
2
3  SELECT DISTINCT ?name
4  WHERE {
5      ?person foaf:name ?name .
6  }
```


SELECT *

```
1  PREFIX foaf: <http://xmlns.com/foaf/0.1/>
2
3  SELECT *
4  WHERE {
5      ?person foaf:name ?name .
6      ?person foaf:mbox ?email .
7  }
8
```

FROM

```
1 PREFIX foaf: <http://xmlns.com/foaf/0.1/>
2 PREFIX card: <http://www.w3.org/People/Berners-Lee/card#>
3
4 SELECT ?homepage
5 FROM <http://dig.csail.mit.edu/2008/webdav/timbl/foaf.rdf>
6 WHERE {
7     card:i foaf:knows ?known .
8     ?known foaf:homepage ?homepage .
9 }
10
```

FILTER

```
1 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
2 PREFIX type: <http://dbpedia.org/class/yago/>
3 PREFIX prop: <http://dbpedia.org/property/>
4
5 SELECT ?country_name ?population
6 WHERE {
7     ?country a type:LandlockedCountries ;
8         rdfs:label ?country_name ;
9         prop:populationEstimate ?population .
10    FILTER (?population > 15000000) .
11 }
12
```

Оператори

Оператор	Опис
!, &&,	Логички оператори
+, -, *, /	Аритметички оператори
=, !=, >, <, IN, NOT IN	Релациони оператори
isURI, isBlank, isLiteral, isNumeric, bound	SPARQL тестови
URI, BNODE, STRDT, STRLANG, UUID, STRUUID	SPARQL конструктори

Table 2: Оператори

Функције

Функција	Опис
STRLEN, SUBSTR, UCASE, LCASE, STRSTARTS, STREND, CONTAINS, CONCAT, itd.	string функције
abs, round, ceil, floor, itd.	аритметичке функције
now, year, month, day, hours, minutes, seconds, timezone, tz	date/time функције
MD5, SHA1, SHA256, SHA384, SHA512	hash функције

Table 3: Функције

OPTIONAL

```
1 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
2 PREFIX db: <http://bio2rdf.org/drugbank_vocabulary:>
3
4 SELECT ?drug_name ?dosage ?indication
5 WHERE {
6     ?drug a db:Drug .
7     ?drug rdfs:label ?drug_name .
8     OPTIONAL {
9         ?drug db:dosage ?dosage .
10    }
11    OPTIONAL {
12        ?drug db:indication ?indication .
13    }
14 }
15
```

UNION

```
1 PREFIX go: <http://purl.org/obo/owl/GO#>
2 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
3 PREFIX obo: <http://www.obofoundry.org/ro/ro.owl#>
4
5 SELECT DISTINCT ?label ?process
6 WHERE {
7     {
8         ?process obo:part_of go:GO_0007165
9     }
10    UNION
11    {
12        ?process rdfs:subClassOf go:GO_0007165
13    }
14    ?process rdfs:label ?label
15 }
16
```

ORDER BY

```
1 PREFIX foaf: <http://xmlns.com/foaf/0.1/>
2
3 SELECT ?name
4 WHERE {
5     ?x foaf:name ?name
6 }
7 ORDER BY ?name
8
```


LIMIT/OFFSET

```
1  PREFIX foaf: <http://xmlns.com/foaf/0.1/>
2
3  SELECT ?name
4  WHERE {
5      ?x foaf:name ?name
6  }
7  ORDER BY ?name
8  LIMIT 5
9  OFFSET 10
10
```

GROUP BY

```
1 PREFIX : <http://example.com/data/#>
2
3 SELECT ?x (MIN(?y) * 2 AS ?min)
4 WHERE {
5     ?x :p ?y .
6 }
7 GROUP BY ?x
8
```

HAVING

```
1  PREFIX : <http://data.example/>
2
3  SELECT (AVG(?size) AS ?asize)
4  WHERE {
5      ?x :size ?size
6  }
7  GROUP BY ?x
8  HAVING(AVG(?size) > 10)
9
```

FROM NAMED

- SPARQL упити се извршавају над RDF скуповима података (тј. над RDF графовима)
- Претходни SPARQL упити су се извршавали над једним (подразумеваним) графом
- RDF скупови података састоје се од подразумеваног графа и нула или више именованих графова (идентификованих URI)

FROM NAMED

- Именовани графови специфицирају се са FROM NAMED клаузулом
- GRAPH клаузула се користи да би се делови упита покlopили са неким именованим графом у RDF скупу података (делови упита изван GRAPH клаузуле се поклапају са подразумеваним графом)

FROM NAMED

```
1  PREFIX foaf: <http://xmlns.com/foaf/0.1/>
2
3  SELECT DISTINCT ?name
4  FROM <http://example.org/alice>
5  FROM NAMED <http://example.org/bob>
6  WHERE {
7      ?person foaf:name ?name .
8      GRAPH <http://example.org/bob> { ?person a foaf:Person } .
9  }
10
```

Резултати SPARQL упита

Резултати SPARQL упита могу да се серијализују у различитим форматима:

- XML (SPARQL Query Results XML Format)
- JSON (SPARQL 1.1 Query Results JSON Format)
- CSV/TSV (SPARQL 1.1 Query Results CSV and TSV Formats)
- RDF (може се серијализовати у више PDF формата)
- HTML (када се са SPARQL упитима ради у интерактивном моду)

CONSTRUCT упит

- CONSTRUCT упит је алтернатива SELECT упиту која враћа (RDF) граф уместо релације (табеле)
- RDF граф се конструише тако што се вредности променљивих у обрасцима графа у CONSTRUCT клаузули мењају резултатима еквивалентног SELECT упита

CONSTRUCT упит

```
1  PREFIX vCard: <http://www.w3.org/2001/vcard-rdf/3.0#>
2  PREFIX foaf: <http://xmlns.com/foaf/0.1/>
3
4  CONSTRUCT {
5      ?X vCard:FN ?name .
6      ?X vCard:URL ?url .
7      ?X vCard:TITLE ?title .
8  }
9  FROM <http://dig.csail.mit.edu/2008/webdav/timbl/foaf.rdf>
10 WHERE {
11     ?X foaf:name ?name .
12     ?X foaf:homepage ?url .
13     ?X foaf:title ?title .
14 }
15
```

ASK упит

- ASK упит враћа boolean вредност у зависности од тога да ли се образац графа поклапа са неким (било којим) скупом података

ASK упит

```
1 PREFIX prop: <http://dbpedia.org/property/>
2
3 ASK
4 {
5     <http://dbpedia.org/resource/Amazon_River> prop:length ?amazon .
6     <http://dbpedia.org/resource/Nile> prop:length ?nile .
7     FILTER(?amazon > ?nile) .
8 }
9
```

DESCRIBE упит

- DESCRIBE упит враћа RDF граф који описује задати ресурс
- Нису интероперабилни зато што различити сервери могу да их интерпретирају на различите начине

DESCRIBE упит

```
1 PREFIX foaf: <http://xmlns.com/foaf/0.1/>
2
3 DESCRIBE ?drug
4 WHERE {
5   ?drug <http://bio2rdf.org/drugbank_vocabulary:brand> "Botox" .
6 }
7
```

Преглед садржаја

- 1 Увод
- 2 Resource Description Framework
- 3 SPARQL
- 4 **Закључак**

Закључак

- Повезани подаци
- Принципи повезаних података
- RDF
- RDF граф
- RDF исказ
- Елементи RDF исказа
- Turtle

Закључак

- SPARQL
- Врсте SPARQL упита
- SPARQL клаузуле
- Образац графа

Литература

- Wood, D., Zaidman, M., Ruth, L., Hausenblas, M. (2014) "Linked Data: Structured Data on the Web". Shelter Island, NY, USA: Manning Publications.

Литература

- World Wide Web Consortium (W3C), Linked Data Platform 1.0, <https://www.w3.org/TR/ldp/>
- World Wide Web Consortium (W3C), RDF 1.1 Concepts and Abstract Syntax, <https://www.w3.org/TR/rdf11-concepts/>
- World Wide Web Consortium (W3C), RDF 1.1 Semantics, <https://www.w3.org/TR/rdf11-mt/>
- World Wide Web Consortium (W3C), RDF 1.1 Primer, <https://www.w3.org/TR/rdf11-primer/>
- World Wide Web Consortium (W3C), RDF 1.1 Turtle, <https://www.w3.org/TR/turtle/>

Хвала на пажњи!