

LAB 1 - Introduction to the Semantic Web

1 Extracting RDF metadata

```
1: <rdf:RDF
2:   xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
3:   xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
4:   xmlns:dc="http://purl.org/dc/elements/1.1/"
5:   xmlns:foaf="http://xmlns.com/foaf/0.1/"
6:   xmlns:image="http://jibbering.com/vocabs/image/#"
7:   xmlns:an="http://www.w3.org/2000/10/annotation-ns#"
8: >
9: <foaf:Image rdf:about="http://www.kanzaki.com/works/2004/imgdsc/miniduck_by_ccc.jpg">
10:  <image:hasPart>
11:    <image:Rectangle rdf:ID="p1">
12:      <image:points>155,63 244,139</image:points>
13:      <dc:title>głowa 1</dc:title>
14:      <dc:description>Głowa kaczki pierwszej</dc:description>
15:      <image:depicts rdf:parseType="Resource">
16:        <dc:description>Głowa kaczki</dc:description>
17:      </image:depicts>
18:    </image:Rectangle>
19:  </image:hasPart>
20:  <image:hasPart>
21:    <image:Rectangle rdf:ID="p2">
22:      <image:points>60,38 138,93</image:points>
23:      <dc:title>głowa 2</dc:title>
24:      <dc:description>Głowa kaczki drugiej</dc:description>
25:      <image:depicts rdf:parseType="Resource">
26:        <dc:description>Głowa kaczki</dc:description>
27:      </image:depicts>
28:    </image:Rectangle>
29:  </image:hasPart>
30:  <image:hasPart>
31:    <image:Rectangle rdf:ID="p3">
32:      <image:points>145,23 186,57</image:points>
33:      <dc:title>dziób kaczki</dc:title>
34:      <dc:description>Dziób kaczki trzeciej</dc:description>
35:      <image:depicts rdf:parseType="Resource">
36:        <dc:description>Dziób kaczki</dc:description>
```

37: </image:depicts>
 38: </image:Rectangle>
 39: </image:hasPart>
 40: </foaf:Image>
 41: </rdf:RDF>

Number	Subject	Predicate	Object
1	http://www.kanzaki.com/works/2004/imgdsc/miniduck_by_ccc.jpg	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://xmlns.com/foaf/0.1/Image
2	http://www.w3.org/RDF/Validator/run/1336211928250#p1	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://jibbering.com/vocabs/image/#Rectangle
3	http://www.kanzaki.com/works/2004/imgdsc/miniduck_by_ccc.jpg	http://jibbering.com/vocabs/image/#hasPart	http://www.w3.org/RDF/Validator/run/1336211928250#p1
4	http://www.w3.org/RDF/Validator/run/1336211928250#p1	http://jibbering.com/vocabs/image/#points	"155,63 244,139"
5	http://www.w3.org/RDF/Validator/run/1336211928250#p1	http://purl.org/dc/elements/1.1/title	"głowa 1"
6	http://www.w3.org/RDF/Validator/run/1336211928250#p1	http://purl.org/dc/elements/1.1/description	"Głowa kaczki pierwszej"
7	http://www.w3.org/RDF/Validator/run/1336211928250#p1	http://jibbering.com/vocabs/image/#depicts	genid:A68529
8	genid:A68529	http://purl.org/dc/elements/1.1/description	"Głowa kaczki"
9	http://www.w3.org/RDF/Validator/run/1336211928250#p2	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://jibbering.com/vocabs/image/#Rectangle
10	http://www.kanzaki.com/works/2004/imgdsc/miniduck_by_ccc.jpg	http://jibbering.com/vocabs/image/#hasPart	http://www.w3.org/RDF/Validator/run/1336211928250#p2
11	http://www.w3.org/	http://jibbering.com/	"60,38 138,93"

	RDF/Validator/run/1336211928250#p2	vocabs/image/#points	
12	http://www.w3.org/RDF/Validator/run/1336211928250#p2	http://purl.org/dc/elements/1.1/title	"głowa 2"
13	http://www.w3.org/RDF/Validator/run/1336211928250#p2	http://purl.org/dc/elements/1.1/description	"Głowa kaczki drugiej"
14	http://www.w3.org/RDF/Validator/run/1336211928250#p2	http://jibbering.com/vocabs/image/#depicts	genid:A68530
15	genid:A68530	http://purl.org/dc/elements/1.1/description	"Głowa kaczki"
16	http://www.w3.org/RDF/Validator/run/1336211928250#p3	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://jibbering.com/vocabs/image/#Rectangle
17	http://www.kanzaki.com/works/2004/imgdsc/miniduck_by_ccc.jpg	http://jibbering.com/vocabs/image/#hasPart	http://www.w3.org/RDF/Validator/run/1336211928250#p3
18	http://www.w3.org/RDF/Validator/run/1336211928250#p3	http://jibbering.com/vocabs/image/#points	"145,23 186,57"
19	http://www.w3.org/RDF/Validator/run/1336211928250#p3	http://purl.org/dc/elements/1.1/title	"dziób kaczki"
20	http://www.w3.org/RDF/Validator/run/1336211928250#p3	http://purl.org/dc/elements/1.1/description	"Dziób kaczki trzeciej"
21	http://www.w3.org/RDF/Validator/run/1336211928250#p3	http://jibbering.com/vocabs/image/#depicts	genid:A68531
22	genid:A68531	http://purl.org/dc/elements/1.1/description	"Dziób kaczki"



2 Common Vocabularies

SKOS - tezaury, schematy klasyfikacji, taksonomie, folksonomie;

Użycie: Biblioteka Kongresu USA, reprezentacja typów danych geograficznych w ontologii GeoNames

Dublin Core - opis wszelkiego rodzaju zasobów: internetowych (audio, wideo, www) czy fizycznych (książki, dzieła sztuki, itp.);

Użycie: OMF (Open Source Metadata Framework), pośrednio systemy pomocy GNOME i KDE

FOAF - opis osób oraz aktywności, w których uczestniczą, a także powiązań pomiędzy różnymi ludźmi

Użycie: serwisy bloggerskie Live Journal i Dead Journal, sieć społecznościowa MyOpera, Wordpress, Yandey

DOAP - opis projektów programistycznych, zwłaszcza open-source;

Użycie: Freshmeat - serwis z oprogramowaniem open-source

3 FOAF

<rdf:RDF

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

```

xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
xmlns:foaf="http://xmlns.com/foaf/0.1/"
xmlns:admin="http://webns.net/mvcb/">
<foaf:PersonalProfileDocument rdf:about="">
  <foaf:maker rdf:resource="#me"/>
  <foaf:primaryTopic rdf:resource="#me"/>
  <admin:generatorAgent rdf:resource="http://www.ldodds.com/foaf/foaf-a-matic"/>
  <admin:errorReportsTo rdf:resource="mailto:leigh@ldodds.com"/>
</foaf:PersonalProfileDocument>
<foaf:Person rdf:ID="me">
<foaf:name>Bartłomiej Hyży</foaf:name>
<foaf:title>Mr</foaf:title>
<foaf:givenname>Bartłomiej</foaf:givenname>
<foaf:family_name>Hyży</foaf:family_name>
<foaf:nick>emp</foaf:nick>
<foaf:mbox_sha1sum>75d3a52c145c9515fb37ec9422e69b74f042b990</foaf:mbox_sha1sum>
<foaf:homepage rdf:resource="http://bartlomiej-hyzy.pl"/>
<foaf:phone rdf:resource="tel:12345678"/>
<foaf:schoolHomepage rdf:resource="http://agh.edu.pl"/>
<foaf:knows>
<foaf:Person>
<foaf:name>Witold Baran</foaf:name>
<foaf:mbox_sha1sum>0efca6b0f5a64e8fc5e96e6c18a6cdf5b681108a</foaf:mbox_sha1sum></
foaf:Person></foaf:knows>
<foaf:knows>
<foaf:Person>
<foaf:name>Adam Małysz</foaf:name>
<foaf:mbox_sha1sum>6a83a836d342aacabf77d17c7713ca54959efca5</
foaf:mbox_sha1sum></foaf:Person></foaf:knows>
<foaf:knows>
<foaf:Person>
<foaf:name>John Travolta</foaf:name>
<foaf:mbox_sha1sum>f571f96c4effd7b6e4f1962b8100ee436503c20d</foaf:mbox_sha1sum></
foaf:Person></foaf:knows></foaf:Person>
</rdf:RDF>

```

FOAF URL: <http://student.agh.edu.pl/~emp/foaf.rdf>

4 Schemas

Ontologie zawarte w bazie <http://www.daml.org/ontologies/> powstawały w latach 2000-2003. Znajdują się tam ontologie o różnych wielkościach - od bardzo prostych, zawierających zaledwie 1-2 elementy, to bardzo duże, składające się z kilkunastu tysięcy elementów.

5 Ontology visualisation

6 Ontology reasoning

Niestety tych dwóch zadań nie zdołaliśmy zrealizować z powodu niefunkcjonowania aplikacji OwlSight.

7 Linked Open Data

DBpedia - ustrukturyzowane informacje pochodzące z Wikipedii i powiązania pomiędzy nimi

ACM - dane opisujące publikacje Association for Computing Machinery

flickr wrappr - zdjęcia pochodzące z serwisu flickr, ilustrujące hasła z DBpedii (Wikipedii)

Linked MDB - informacje o filmach i powiązania między nimi

Project Gutenberg (FUB) - rekordy opisujące książki elektroniczne dostępne za pośrednictwem Projektu Gutenberg

8 Semantic Web tools

RDFizers - według naszej opinii naistotniejszą są te, które przekształcają treści o charakterze edukacyjnym. Z tego względu konwersje javy oraz javadoca wydały nam się najciekawsze. Inne konwersje, które byśmy dodali to na przykład konwersja mp3 ID3 => RDF.

Semantic Web browser - na danej stronie przedstawiono 13 przeglądarek, wśród których dwa przekierowania są nieaktualne.

Semantic Web Search Engines - są bardzo proste w użyciu. Wystarczy wpisać szukany tekst w odpowiednie pole, a silnika zajmie się przeszukiwaniem dokumentów RDF. Zostają wyświetlone szczegółowe informacje dotyczące wyników wyszukiwania.

Applications of the Linked Data - bardzo użyteczna może być DBpedia, która daje możliwość tworzenia skomplikowanych zapytań złożonych z relacji i zależności pomiędzy danymi, np. "znajdz drużyny ktore przynajmniej 3 razy wygrał mistrzostwo świata w piłce noznej"

Control questions

- How one can add semantic annotation to a web page?

Można ręcznie dodać adnotacje RDF strony. Niesie to za sobą możliwość popełnienia błędów, przez co sieć semantyczna nie będzie działać prawidłowo. Dlatego też czasem lepiej jest wykorzystać gotowe aplikacje które generują dane RDF (np. flickr2rdf).

- **Explain the Semantic Web Stack of technologies.**

Semantic Web Stack of technologies przedstawia hierarchię języków i technologii zorganizowanych w formie warstw, wykorzystywane w tworzeniu sieci semantycznej. Każda warstwa wyższa korzysta z funkcji warstwy niższej.

Warstwy:

URI/IRI - identyfikacja zasobu sieci semantycznej

XML layer - podstawy składni sieci semantycznej

RDF layer - podstawowy model danych dla faktów, proste schematy ontologii

Ontology layer - bardziej zaawansowane ontologie niż schematy RDF (standard OWL)

Logic layer - rozwinięcie ontologii, analiza wiedzy specyficznej dla danej aplikacji

Proof layer - walidacja i wymiana danych

Trust layer - weryfikacja czy dane pochodzą z zaufanego źródła

User Interface & applications - umożliwianie użytkownikom korzystanie z sieci semantycznej

-

Główną składnią RDF są trójki (ang. triples) składające się z:

1. podmiotu ("książka")
2. orzeczenia/predykatu ("o tytule")
3. dopełnienia/obiektu ("Alicja w krainie czarów")

Zaletą tego formatu w stosunku do innych składni jest jego prostota oraz niezmienność. RDF pozwala na opisanie w ten sam, ustandaryzowany sposób różnych rodzajów informacji.

- **What is the role of the ontologies in the Semantic Web?**

Rolą ontologii w sieciach semantycznych jest określanie znaczenia zasobów informacyjnych. Definiują zależności między danymi, dzięki czemu sieci mogą przeprowadzać wnioskowanie.

- **What are the 4 principles of Linked Data?**

1. Używanie URI do opisywania "rzeczy"
2. Używanie URI HTTP, dzięki czemu możliwe jest ich użycie przez ludzi i maszyny
3. Używanie standardowych formatów (np. RDF/XML) do opisu "rzeczy" po dereferencji ich URI
4. Umieszczanie w opisie "rzeczy" powiązań i relacji z innymi obiektami i informacjami (powiązane dane semantyczne)

LAB 2 - eXtensible Markup Language (XML)

1 Warm-up: Well-formedness of the XML

- `<a><a>`

Niepoprawny: brak zamknięcia elementu `<a>`

- `<a>foo<a>bar`

Niepoprawny: brak elementu nadrzędnego dokumentu (root element)

- `<a>foobar`

Niepoprawny: brak zamknięcia elementu `bar`

- `<a><c>foobar</c>`

Niepoprawny: nieprawidłowe zagnieżdżenie elementów (zła kolejność zamknięcia elementów `` i `<c>`)

- `<a/>foobar`

Niepoprawny: brak elementu nadrzędnego dokumentu, nie zamknięty ostatni element ``

- `<a><c>foo</c>baz<c>bar</c>`

Poprawny:

```
a
--b
----c
-----"foo"
----"baz"
----c
-----"bar"
```

- `foo`

Poprawny:

```
a (x="1",X="2")
--b
----"foo"
```

- `<b x="2">foo`

Poprawny:

```
a (x="1")
--b (x="2")
----"foo"
```

- `foo`

Niepoprawny: podwójnie zdefiniowany atrybut "x" elementu `<a>`


```
- <a><b x="3"><c y="1">foo</c><c>bar</c></b></a>
```

Poprawny:

a

```
--b (x="3")
```

```
----c (y="1")
```

```
-----"foo"
```

```
----c
```

```
-----"bar"
```

2 Creating XML document

```
<book asin="B000LNTQTM">
  <title>Compilers: Principles, Techniques, and Tools</title>
  <publisher>Addison-Wesley</publisher>
  <year>1985</year>
  <authors>
    <author>Alfred V. Aho</author>
    <author>Ravi Sethi</author>
    <author>Jeffrey D. Ullman</author>
  </authors>
</book>
```

3 DTD

```
<!DOCTYPE book [
  <!ELEMENT book (title,publisher,year,authors)>
  <!ELEMENT title (#PCDATA)>
  <!ELEMENT publisher (#PCDATA)>
  <!ELEMENT year (#PCDATA)>
  <!ELEMENT authors (author+)>
  <!ELEMENT author (#PCDATA)>
  <!ATTLIST book asin CDATA #REQUIRED>
]>
```

5 XSD

XML Schema:

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

```

<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema">
  <xs:element name="book">
    <xs:complexType>
      <xs:sequence>
        <xs:element name="title" type="xs:string"/>
        <xs:element name="publisher" type="xs:string"/>
        <xs:element name="year" type="xs:positiveInteger"/>
        <xs:element name="authors">
          <xs:complexType>
            <xs:sequence>
              <xs:element name="author" type="xs:string"
                minOccurs="1" maxOccurs="unbounded"/>
            </xs:sequence>
          </xs:complexType>
        </xs:element>
      </xs:sequence>
      <xs:attribute name="asin" type="xs:string" use="required"/>
    </xs:complexType>
  </xs:element>
</xs:schema>

```

Testowy dokument XML:

```

<?xml version="1.0" encoding="utf-8"?>
<book asin="B000LNTQTM"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:noNamespaceSchemaLocation="book.xsd">
  <title>Compilers: Principles, Techniques, and Tools</title>
  <publisher>Addison-Wesley</publisher>
  <year>1985</year>
  <authors>
    <author>Alfred V. Aho</author>
    <author>Ravi Sethi</author>
    <author>Jeffrey D. Ullman</author>
  </authors>
</book>

```