

**Akademia Górniczo-Hutnicza
im. Stanisława Staszica w Krakowie**



**Wydział Elektrotechniki, Automatyki, Informatyki i
Elektroniki**

**Podstawy Inżynierii Wiedzy
Semantic Web**

**Sprawozdanie:
Introduction to the Semantic Web
XML**

**Wykonali:
Paweł Maślanka
Norbert Pabian**

I Introduction to the Semantic Web

1 Extracting RDF metadata

```
<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:dc="http://purl.org/dc/elements/1.1/"
  xmlns:foaf="http://xmlns.com/foaf/0.1/"
  xmlns:image="http://jibbering.com/vocabs/image/#"
  xmlns:an="http://www.w3.org/2000/10/annotation-ns#"
>
  <foaf:Image
rdf:about="http://www.kanzaki.com/works/2004/imgdsc/miniduck_by_ccc.jpg">
  <image:hasPart>
    <image:Rectangle rdf:ID="p1">
      <image:points>183,78 243,133</image:points>
      <dc:title>Głowa kaczk</dc:title>
      <dc:description>Głowa gumowej kaczk</dc:description>
      <image:depicts rdf:parseType="Resource">
        <dc:description>Przedstawia głowę kaczk</dc:description>
        </image:depicts>
      </image:Rectangle>
    </image:hasPart>
  </foaf:Image>
</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/1336505958715#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/1336505958715#p1
4	http://www.w3.org/RDF/Validator/run/1336505958715#p1	http://jibbering.com/vocabs/image/#points	"183,78 243,133"
5	http://www.w3.org/RDF/Validator/run/1336505958715#p1	http://purl.org/dc/elements/1.1/title	"Głowa kaczk"
6	http://www.w3.org/RDF/Validator/run/1336505958715#p1	http://purl.org/dc/elements/1.1/description	"Głowa gumowej kaczk"
7	http://www.w3.org/RDF/Validator/run/1336505958715#p1	http://jibbering.com/vocabs/image/#depicts	genid:A50203
8	genid:A50203	http://purl.org/dc/elements/1.1/description	"Przedstawia głowę kaczk z najbliższego planu"



```
<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:dc="http://purl.org/dc/elements/1.1/"
  xmlns:foaf="http://xmlns.com/foaf/0.1/"
  xmlns:image="http://jibbering.com/vocabs/image/#"
  xmlns:an="http://www.w3.org/2000/10/annotation-ns#"
>
  <foaf:Image
rdf:about="http://www.kanzaki.com/works/2004/imgdsc/miniduck_by_ccc.jpg">
  <image:hasPart>
    <image:Rectangle rdf:ID="p1">
      <image:points>183,78 243,133</image:points>
      <dc:title>Głowa kaczk</dc:title>
      <dc:description>Głowa najbliższej kaczk</dc:description>
      <image:depicts rdf:parseType="Resource">
        <dc:description>Przedstawia głow kaczk</dc:description>
      </image:depicts>
    </image:Rectangle>
  </image:hasPart>
  <image:hasPart>
    <image:Rectangle rdf:ID="p2">
      <image:points>31,14 281,218</image:points>
      <dc:title>Trzy kaczk</dc:title>
      <dc:description>Trzy gumowe kaczk</dc:description>
      <image:depicts rdf:parseType="Resource">
        <dc:description>Przedstawia na jednym planie trzy gumowe
          kaczk</dc:description>
      </image:depicts>
    </image:Rectangle>
  </image:hasPart>
</foaf:Image>
</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/1336508892065#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/1336508892065#p1
4	http://www.w3.org/RDF/Validator/run/1336508892065#p1	http://jibbering.com/vocabs/image/#points	"183,78 243,133"
5	http://www.w3.org/RDF/Validator/run/1336508892065#p1	http://purl.org/dc/elements/1.1/title	"Głowa kaczk"

	92065#p1		
6	http://www.w3.org/RDF/Validator/run/1336508892065#p1	http://purl.org/dc/elements/1.1/description	"Głowa najbliższej kaczki"
7	http://www.w3.org/RDF/Validator/run/1336508892065#p1	http://jibbering.com/vocabs/image/#dependencies	genid:A50339
8	genid:A50339	http://purl.org/dc/elements/1.1/description	"Przedstawia głowę kaczki"
9	http://www.w3.org/RDF/Validator/run/1336508892065#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/1336508892065#p2
11	http://www.w3.org/RDF/Validator/run/1336508892065#p2	http://jibbering.com/vocabs/image/#points	"31,14 281,218"
12	http://www.w3.org/RDF/Validator/run/1336508892065#p2	http://purl.org/dc/elements/1.1/title	"Trzy kaczki"
13	http://www.w3.org/RDF/Validator/run/1336508892065#p2	http://purl.org/dc/elements/1.1/description	"Trzy gumowe kaczki"
14	http://www.w3.org/RDF/Validator/run/1336508892065#p2	http://jibbering.com/vocabs/image/#dependencies	genid:A50340
15	genid:A50340	http://purl.org/dc/elements/1.1/description	"Przedstawia na jednym planie trzy gumowe kaczki"



2 Common Vocabularies

2.2.I What is each vocabulary for?

- **SKOS** - jest to rodzina słowników służąca do dokumentowania taksonomi, tezaursów czy schematów klasyfikacyjnych, ich łączenia z innymi danymi w celu publikacji w sieci
- **Dublin Core** - standard metadanych, służący do opisu zasobów internetowych z wykorzystaniem 15 zdefiniowanych elementów
- **FOAF** - jest formatem służącym do opisu osób, dzięki czemu opisy mogą być przetwarzane przez maszyny, dzięki możliwościom odwołań do innych dokumentów tworzą sieci powiązań
- **DOAP** - format służący do opisu projektów informatycznych, w szczególności Open Source.

2.2.II What application use them?

- **SKOS:**
 - Library of Congress Subject Headings
 - UN FAO project AGROVOC
- **Dublin Core:**
 - Open Source Metadata Framework
 - PBCore
 - Zope CMF's
 - ERP5
- **FOAF:**
 - Semantic Web Applications
 - Wordpress
 - Yandex
 - My Opera
- **DOAP:**
 - Apache Software Foundation
 - Python Package Index
 - Dolibarr ERP/CRM

3 FOAF

Paweł:

```
<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>Paweł Maślanka</foaf:name>
<foaf:title>Mr</foaf:title>
<foaf:givenname>Paweł</foaf:givenname>
<foaf:family_name>Maślanka</foaf:family_name>
<foaf:nick>pawmas</foaf:nick>
<foaf:mbox_sha1sum>6cd534902fbdf8201769510768472d7b79a92215</foaf:mbox_sha1sum>
<foaf:homepage rdf:resource="www.pawmas.pl"/>
<foaf:workplaceHomepage rdf:resource="www.pawmas.pl"/>
<foaf:workInfoHomepage rdf:resource="Programming"/>
<foaf:schoolHomepage rdf:resource="www.agh.edu.pl"/>
<foaf:knows>
```

```
<foaf:Person>
<foaf:name>Piotrek</foaf:name>
<foaf:mbox_sha1sum>e199c4554f5d88943b9d5a523f7890e7ff44944b</foaf:mbox_sha1sum></foaf:Person></foaf:Person></foaf:knows></foaf:Person>
</rdf:RDF>
```

URL: http://www.pmj.is.net.pl/semanticweb/pawel_foaf.rdf

Norbert:

```
<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>Norbert Pabian</foaf:name>
<foaf:title>Mr</foaf:title>
<foaf:givenname>Norbert </foaf:givenname>
<foaf:family_name>Pabian</foaf:family_name>
<foaf:nick>fuko fabio</foaf:nick>
<foaf:mbox_sha1sum>216438716ff62403ebd87c2c6be6835b34dd8351</foaf:mbox_sha1sum>
<foaf:schoolHomepage rdf:resource="www.agh.edu.pl"/></foaf:Person>
</rdf:RDF>
```

URL: http://www.pmj.is.net.pl/semanticweb/norbert_foaf.rdf

4 Schemas

Ontologie pojawiały się w latach 2000 - 2003. Ich rozmiar jest różny, zaczynając od małych i prostych 1-2 elementowych ontologii, kończąc na wielkich, złożonych ontologiach, składających się z kilku tysięcy elementów.

5 Ontology visualisation

Wykonanie tego punktu było niemożliwe z powodu błędu serwera podanej aplikacji: **Service Temporarily Unavailable**

6 Ontology reasoning

Wykonanie tego punktu było niemożliwe z powodu błędu serwera podanej aplikacji: **Service Temporarily Unavailable**

7 Linked Open Data

- **DBpedia** - zawiera dane z wikipedii złożone w struktury wygodne do odczytu
- **GeoNames** - ontologie opisują i wiążą ze sobą regiony, nazwy geograficzne, a także wszystkie inne informacje które są możliwe do znalezienia na typowej mapie.
- **IEEE** – zawiera dane z artykułów **IEEE**
- **Open Library** - zawiera dane wielu milionów książek. Celem projektu jest stworzeniu spisu wszystkich opublikowanych na świecie dzieł literackich

- **Medi Care** – lista leków oraz dane na ich temat

8 Semantic Web tools

- **RDFizers**

Najciekawsze wydają się być rdfizery, które pozwalają budować reprezentacje RDF dla danych które są bardzo często przez nas wykorzystywane. Dlatego też do najbardziej przydatnych można by zaliczyć: JPEG -> RDF, EMail -> RDF, Bibtex -> RDF, Java -> RDF .

Jeśli chodzi o wymyślenie nowych, to interesujące byłoby dysponowanie rdfizerem, który byłby zdolny na konwersję plików graficznych, popularnych edytorów grafiki, np. PSD → RDF.

- **Semantic Web browsers**

Przegląd linków wykazał, że spośród podanych działa lub jest możliwych do instalacji 8 wyszukiwarek semantycznych sieci. Nie załadowały się: razorbase, iLOD, FenFire, Objectviewer, zLinks + Tabulator działa tylko na Firefoxie i Camino.

- **Semantic Web Search Engines**

Łatwo z nich korzystać. Użytkownik musi jedynie umieścić frazę, jaką chce wyszukać w odpowiednim polu, a następnie kliknąć przycisk **Wyszukaj**. Silnik wyszukiwarki zwraca precyzyjne wyniki w plikach RDF.

- **Applications of the Linked Data**

Najciekawszymi wydają się być te, które będą miały szerokie zastosowanie. **LinkedGeoData** Browser pozwala na wygodne przeglądanie danych geograficznych, a także ich wyszukiwanie. Tabulator pozwala na wygodne przeglądanie danych w formacie RDF, jest on dostępny także w formie wtyczki do przeglądarki Firefox, co pozwala na wygodne przeglądanie różnego rodzaju „linked data”.

Control questions

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

Wystarczy, że doda się do kodu strony odpowiednią strukturę zgodną z regułami danego formatu np. parametr rel w tagu <a> (**XFN - XHTML Friends Network**)

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

Jest to stos przedstawiający architekturę Semantic Web, poczynając od interfejsu użytkownika, poprzez warstwy ontologii i taksonomii, składnię aż po URI i zbiór znaków.

- **What is the main syntax for RDF? What are its advantages over other syntaxes?**

RDF zawiera następujące klasy:

- rdf:XMLLiteral
- rdf:Property
- rdf:Statement
- rdf:Alt, rdf:Bag, rdf:Seq
- rdf:List
- rdf:nil

oraz następujące właściwości:

- rdf:type
- rdf:Property
- rdf:first
- rdf:rest

- rdf:value
- rdf:subject
- rdf:predicate
- rdf:object

Podstawową składnią, określającą ogólne zasady pisania dokumentów RDF, jest język XML. Składnia jest prosta do zrozumienia i wykorzystania przez zwykłego użytkownika Internetu. Umożliwia proste śledzenie oraz wyszukiwanie danych.

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

Ontologie definiują zbiór pojęć dla pewnej dziedziny wiedzy, na której budowana jest Semantic Web. Mogą służyć jako podstawa do wnioskowania właściwości opisywanych przez ontologię pojęć.

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

- Używanie **URI** do identyfikacji rzeczy
- Używanie **HTTP URI**, dzięki czemu będzie można się do tych rzeczy odnosić oraz będą mogły być wyszukiwane przez ludzi oraz aplikacje klienckie.
- Dostarczanie użytecznych informacji o danej rzeczy, wykorzystując takie formaty jak **RDF/XML**, widoczne gdy ktoś odwołuje się do URI.
- Zawieranie odnośników do innych, powiązanych **URI** w celu zwiększenia szansy na znalezienie powiązanych danych w sieci.

II eXtensible Markup Language (XML)

Well-formedness of the XML

Determine which of the following XML documents is well-formed. If a document is well-formed, show its corresponding tree model. If it is not well-formed, explain why not. You may assume that all of these documents are intended to be XML 1.0 documents, and thus the declaration is not needed.

Examples of XML code lines:

```
<a><b/><a>
```

Notation is not correct. No closing tag for the root element `<a >` last element should look like `` and the second element `` is also bad and it should not be there.

```
<a><b>foo</b></a><a>bar</a>
```

Notation is not correct. No root element.

```
<a><b>foo</b><b>bar</a>
```

Notation is not correct. No closing tag for the second `` element.

```
<a><b><c>foo</b>bar</c></a>
```

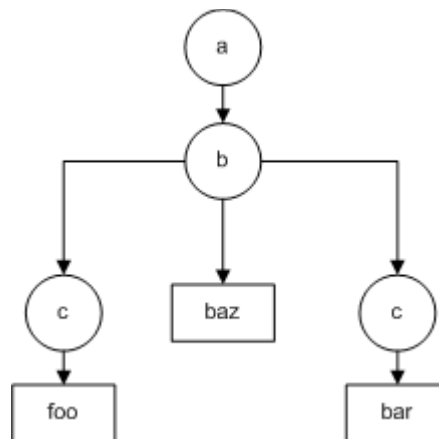
Notation is not correct. Two elements `` and `<c >` have wrong closing tag.

```
<a/><b>foo</b><b>bar<b>
```

Notation is not correct. First element is bad. No root element and last element `` should be closing tag.

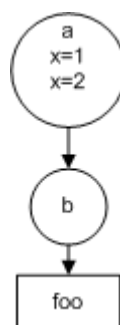
```
<a><b><c>foo</c>baz<c>bar</c></b></a>
```

Notation is well formatted.



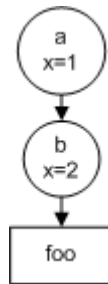
```
<a x="1" X="2"><b>foo</b></a>
```

Notation is well formatted.



```
<a x="1"><b x="2">foo</b></a>
```

Notation is well formatted.

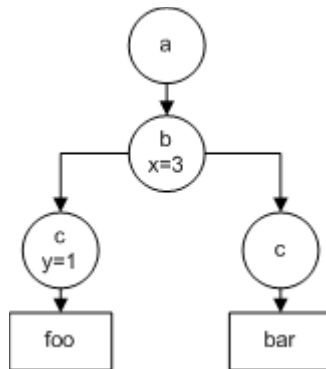


```
<a x="1" y="2" x="3"><b>foo</b></a>
```

Notation is not correct. Parameter x was already defined for < a > element.

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

Notation is well formatted.



Creating XML document

Modify the code below to describe the best book from your bookshelf:

```
<book>  
<author>Grigoris Antoniou</author>  
<author>Frank van Harmelen</author>  
<title>A Semantic Web Primer</title>  
<date>2008-10-09</date>  
<publisher>The MIT Press</publisher>  
</book>
```

My example:

```
<book>  
<author>Adam Mickiewicz</author>  
<title>Pan Tadeusz</title>  
<date>04-07-2010</date>  
<publisher>Siedmioróg, Przedsiębiorstwo Wydawniczo-Handlowe</publisher>  
<type>Poezja</type>  
<pages>424</pages>  
<isbn>83-725-4315-1</isbn>  
</book>
```

Document Type Definition (DTD)

Using a text editor, define a DTD definition for your XML syntax. Put it in the report
You can define DTD definition inside XML file:

```
<?xml version="1.0" ?>
<!DOCTYPE root [
]>
XML content
```

Defined document type:

```
<!DOCTYPE book [
<!ELEMENT book (author, title, date, publisher, type, pages, isbn)>
<!ELEMENT author (#PCDATA)>
<!ELEMENT title (#PCDATA)>
<!ELEMENT date (#PCDATA)>
<!ELEMENT publisher (#PCDATA)>
<!ELEMENT type (#PCDATA)>
<!ELEMENT pages (#PCDATA)>
<!ELEMENT isbn (#PCDATA)>
]>
```

XML Schema (XSD)

Using a text editor, define an XSD definition for your XML syntax.

```
<?xml version="1.0"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema">
<xs:element name="book">
<xs:complexType>
<xs:sequence>
<xs:element name="author" type="xs:string"/>
<xs:element name="title" type="xs:string"/>
<xs:element name="date" type="xs:date"/>
<xs:element name="publisher" type="xs:string"/>
<xs:element name="type" type="xs:string"/>
<xs:element name="pages" type="xs:integer"/>
<xs:element name="isbn" type="xs:string"/>
</xs:sequence>
</xs:complexType>
</xs:element>
</xs:schema>
```