

AKADEMIA GÓRNICZO-HUTNICZA

WYDZIAŁ ELEKTROTECHNIKI, AUTOMATYKI, INFORMATYKI I ELEKTRONIKI
KIERUNEK INFORMATYKA STOSOWANA



PODSTAWY INŻYNIERII WIEDZY

Laboratoria 1 i 2

Introduction to the Semantic Web i XML

Sebastian Gacek, Łukasz Zieńkowski

Kraków, 08 maja 2012

1 Introduction to the Semantic Web

1. Extracting RDF metadata

Wygenerowany kod RDF

```
<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#"
>
  <rdf:Description rdf:about="">
    <an:annotates rdf:resource="http://www.kanzaki.com/works/2004/imgdsc/miniduck_by_ccc.jpg"/>
    <an:created>2012-05-03T10:50:50Z</an:created>
  </rdf:Description>
  <foaf:Image rdf:about="http://www.kanzaki.com/works/2004/imgdsc/miniduck_by_ccc.jpg">
    <dc:title>Kaczki</dc:title>
    <dc:description>Kaczki</dc:description>
    <dc:creator>Creator</dc:creator>
    <dc:date>12-12-12</dc:date>
    <image:width>300</image:width>
    <image:height>225</image:height>
    <image:hasPart>
      <image:Rectangle rdf:ID="p1">
        <image:points>213,96 232,114</image:points>
        <dc:title>0ko prawej kaczki</dc:title>
        <dc:description>Kaczki</dc:description>
        <image:depicts rdf:parseType="Resource">
          <dc:description>Lewe oko kaczki po prawej stronie</dc:description>
        </image:depicts>
      </image:Rectangle>
    </image:hasPart>
    <image:hasPart>
      <image:Rectangle rdf:ID="p2">
        <image:points>58,39 135,93</image:points>
        <dc:title>Głowa lewej kaczki</dc:title>
        <dc:description>Kaczki </dc:description>
        <image:depicts rdf:parseType="Resource">
          <dc:description>Głowa kaczki po lewej stronie</dc:description>
        </image:depicts>
      </image:Rectangle>
    </image:hasPart>
    <image:hasPart>
      <image:Rectangle rdf:ID="p3">
        <image:points>143,32 184,57</image:points>
        <dc:title>Dziór srodkowej kaczki</dc:title>
        <dc:description>Kaczki </dc:description>
        <image:depicts rdf:parseType="Resource">
          <dc:description>Dziób ostatniej kaczki</dc:description>
        </image:depicts>
      </image:Rectangle>
    </image:hasPart>
  </foaf:Image>
</rdf:RDF>
```

Triples of the Data Model

Number	Subject	Predicate	Object
1	http://www.w3.org/RDF/Validator/run/1336296922730	http://www.w3.org/2000/10/annotation-ns#annotates	http://www.kanzaki.com/works/2004/imgdsc/m/iniduck_by_ccc.jpg
2	http://www.w3.org/RDF/Validator/run/1336296922730	http://www.w3.org/2000/10/annotation-ns#created	"2012-05-03T10:50:50Z"
3	http://www.kanzaki.com/works/2004/imgdsc/m/iniduck_by_ccc.jpg	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://xmlns.com/foaf/0.1/Image
4	http://www.kanzaki.com/works/2004/imgdsc/m/iniduck_by_ccc.jpg	http://purl.org/dc/elements/1.1/title	"Kaczki"
5	http://www.kanzaki.com/works/2004/imgdsc/m/iniduck_by_ccc.jpg	http://purl.org/dc/elements/1.1/description	"Kaczki"
6	http://www.kanzaki.com/works/2004/imgdsc/m/iniduck_by_ccc.jpg	http://purl.org/dc/elements/1.1/creator	"Creator"
7	http://www.kanzaki.com/works/2004/imgdsc/m/iniduck_by_ccc.jpg	http://purl.org/dc/elements/1.1/date	"12-12-12"
8	http://www.kanzaki.com/works/2004/imgdsc/m/iniduck_by_ccc.jpg	http://jibbering.com/vocabs/image/#width	"300"
9	http://www.kanzaki.com/works/2004/imgdsc/m/iniduck_by_ccc.jpg	http://jibbering.com/vocabs/image/#height	"225"
10	http://www.w3.org/RDF/Validator/run/1336296922730#p1	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://jibbering.com/vocabs/image/#Rectangle
11	http://www.kanzaki.com/works/2004/imgdsc/m/iniduck_by_ccc.jpg	http://jibbering.com/vocabs/image/#hasPart	http://www.w3.org/RDF/Validator/run/1336296922730#p1
12	http://www.w3.org/RDF/Validator/run/1336296922730#p1	http://jibbering.com/vocabs/image/#points	"213,96 232,114"
13	http://www.w3.org/RDF/Validator/run/1336296922730#p1	http://purl.org/dc/elements/1.1/title	"Oko prawej kaczki"
14	http://www.w3.org/RDF/Validator/run/1336296922730#p1	http://purl.org/dc/elements/1.1/description	"Kaczki"
15	http://www.w3.org/RDF/Validator/run/1336296922730#p1	http://jibbering.com/vocabs/image/#id	genid:A76165

	6922730#p1	epicts	
16	genid:A76165	http://purl.org/dc/elements/1.1/description	"Lewe oko kaczki po prawej stronie"
17	http://www.w3.org/RDF/Validator/run/1336296922730#p2	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://jibbering.com/vocabs/image/#Rectangle
18	http://www.kanzaki.com/works/2004/imgdsc/miniduck_by_ccc.jpg	http://jibbering.com/vocabs/image/#hasPart	http://www.w3.org/RDF/Validator/run/1336296922730#p2
19	http://www.w3.org/RDF/Validator/run/1336296922730#p2	http://jibbering.com/vocabs/image/#points	"58,39 135,93"
20	http://www.w3.org/RDF/Validator/run/1336296922730#p2	http://purl.org/dc/elements/1.1/title	"Głowa lewej kaczki"
21	http://www.w3.org/RDF/Validator/run/1336296922730#p2	http://purl.org/dc/elements/1.1/description	"Kaczki"
22	http://www.w3.org/RDF/Validator/run/1336296922730#p2	http://jibbering.com/vocabs/image/#denotes	genid:A76166
23	genid:A76166	http://purl.org/dc/elements/1.1/description	"Głowa kaczki po lewej stronie"
24	http://www.w3.org/RDF/Validator/run/1336296922730#p3	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://jibbering.com/vocabs/image/#Rectangle
25	http://www.kanzaki.com/works/2004/imgdsc/miniduck_by_ccc.jpg	http://jibbering.com/vocabs/image/#hasPart	http://www.w3.org/RDF/Validator/run/1336296922730#p3
26	http://www.w3.org/RDF/Validator/run/1336296922730#p3	http://jibbering.com/vocabs/image/#points	"143,32 184,57"
27	http://www.w3.org/RDF/Validator/run/1336296922730#p3	http://purl.org/dc/elements/1.1/title	"Dziób środkowej kaczki"
28	http://www.w3.org/RDF/Validator/run/1336296922730#p3	http://purl.org/dc/elements/1.1/description	"Kaczki"
29	http://www.w3.org/RDF/Validator/run/1336296922730#p3	http://jibbering.com/vocabs/image/#denotes	genid:A76167
30	genid:A76167	http://purl.org/dc/elements/1.1/description	"Dziób ostatniej kaczki"

SKOS

Służy do reprezentacji taksonomii, schematów, klasyfikowania itp.
-wykorzystywany przez W3C

Dublin Core

Służy do opisu zasobów internetowych.
-używany w dLibra(biblioteka cyfrowa)

FOAF

Służy do opisu osoby, można tworzyć sieci powizań.
Używany m.in. przez WordPress

DOAP

Służy do opisów projektów(szczególnie typu open-source)
-używany przez Mozilla Foundation

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>Łukasz Zieńkowski</foaf:name>
<foaf:title>Mr</foaf:title>
<foaf:givenname>Łukasz</foaf:givenname>
<foaf:family_name>Zieńkowski</foaf:family_name>
<foaf:nick>Luke</foaf:nick>
<foaf:mbox_sha1sum>6bf1ba3094b06c13f046a3847ba569f090e163e7</foaf:mbox_sha1sum>
<foaf:homepage rdf:resource="imienazwisko.pl"/>
<foaf:depiction rdf:resource="brak"/>
<foaf:phone rdf:resource="tel:555-1234"/>
<foaf:workplaceHomepage rdf:resource="Area 51"/>
<foaf:workInfoHomepage rdf:resource="Secret "/>
<foaf:schoolHomepage rdf:resource="agh.edu.pl"/>
<foaf:knows>
<foaf:Person>
<foaf:name>Jan</foaf:name>
<foaf:mbox_sha1sum>ed47472ea086be0598769020044354f6117b44e3</foaf:mbox_sha1sum>
<rdfs:seeAlso rdf:resource="jan.pl"/></foaf:Person></foaf:knows>
<foaf:knows>
<foaf:Person>
<foaf:name>Jacek</foaf:name>
<foaf:mbox_sha1sum>9a61438ac660d00f694d6a756bbe47163b112d83</foaf:mbox_sha1sum>
<rdfs:seeAlso rdf:resource="jacek.pl"/></foaf:Person></foaf:knows>
<foaf:knows>
<foaf:Person>
<foaf:name>marta</foaf:name>
<foaf:mbox_sha1sum>f8b24484b75af63a738749dddef879572764bda1</foaf:mbox_sha1sum>
<rdfs:seeAlso rdf:resource="marta.pl"/></foaf:Person></foaf:knows>
<foaf:knows>
<foaf:Person>
<foaf:name>monika</foaf:name>
<foaf:mbox_sha1sum>630b487a8a11d8b7447e945018067174deab3f45</foaf:mbox_sha1sum>
<rdfs:seeAlso rdf:resource="monika.pl"/></foaf:Person></foaf:knows>
```

```

<foaf:knows>
<foaf:Person>
<foaf:name>marcin</foaf:name>
<foaf:mbox_sha1sum>190d430090060b81a8aa248f7f1aa764f40d9c2c</foaf:mbox_sha1sum>
<rdfs:seeAlso rdf:resource="marcin.pl"/></foaf:Person></foaf:knows>
<foaf:knows>
<foaf:Person>
<foaf:name>agnieszka</foaf:name>
<foaf:mbox_sha1sum>53876d53328721b6cc688e1245d2b61496be405f</foaf:mbox_sha1sum>
<rdfs:seeAlso rdf:resource="agnieszka.pl"/></foaf:Person></foaf:knows>
<foaf:knows>
<foaf:Person>
<foaf:name>duszan</foaf:name>
<foaf:mbox_sha1sum>7a72cab1339f0dff60cbcb2044837184fcabddd9</foaf:mbox_sha1sum>
<rdfs:seeAlso rdf:resource="duszan.pl"/></foaf:Person></foaf:knows>
<foaf:knows>
<foaf:Person>
<foaf:name>john</foaf:name>
<foaf:mbox_sha1sum>77d0a8b957b9e316620d66631e1254e906148164</foaf:mbox_sha1sum>
<rdfs:seeAlso rdf:resource="john.pl"/></foaf:Person></foaf:knows></foaf:Person>
</rdf:RDF>

```

Adres pliku FOAF <http://student.agh.edu.pl/~dirsot/SemanticWeb/foam.rdf>

4. Schemas

Ontologie pochodzą z lat 2000-2003. Zawierają zarówno małe, jak i bardzo duże ontologie.

5. Ontology visualisation

6. Ontology reasoning

Powyższe dwa punkty nie zostały wykonane z powodu problemów z aplikacją.

7. Linked Open Data

New York Times

Zbiór informacji prasowych, jednego z największych dzienników.

Pokepedia

Zbiór informacji dotyczących pokemonów, serialu animowanego produkcji japońskiej.

DBpedia

Zbiór informacji zawartych z wikipedi w bardzo przystępnej strukturze.

IEEE

Zbiór informacji zawartych w publikacjach wydawnictwa IEEE

ePrints

Zbiór informacji które zostały pobrane eprints.org

8. Semantic Web tools

RDFizers:

Najistotniejsze w mojej opinii:

EMail RDFizer

Maven POM RDFizer

Java RDFizer

Formaty które należałoby uwzględnić:

Semantic Web browsers

Z pośród podanych, dwie przeglądarki są nieaktywne.

Semantic Web Search Engines

Są bardzo intuicyjne w obsłudze. Aktywność użytkownika sprowadza się jedynie do wprowadzenia szukanej frazy

Applications

Naszym zdaniem najistotniejsze są narzędzia typu wiki, dzięki łatwemu systemowi współdzielenia informacji, wiedza może być w szybki i prosty sposób dostarczana do dużej grupy odbiorców.

9. Control questions

How one can add semantic annotation to a web page?

Poprzez edycję kodu strony.

Explain the Semantic Web Stack of technologies.

Hierarchiczna reprezentacja języków i technologii z których składa się aplikacja korzystająca z ontologii. Każda warstwa wyżej korzysta z warstw poniżej jej.

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

Główne elementy to:

-elementy reprezentujące korzeń

-<rdf:RDF>

-<rdf:Description>

Niewątpliwą zaletą takiego syntaksu jest jego prostota, umożliwia ona angażowanie dużej liczby osób, która nie musi posiadać dużej wiedzy, do tworzenia plików w tym formacie.

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

Rolą ontologii jest ułatwienie integracji danych, w wyniku czego wyszukiwanie potrzebnych informacji staje się łatwiejsze.

What are the 4 principles of Linked Data?

Cztery główne zasady to:

-Use URIs to identify things.

-Use HTTP URIs so that these things can be referred to and looked up ("dereferenced") by people and user agents.

-Provide useful information about the thing when its URI is dereferenced, using standard formats such as RDF/XML.

-Include links to other, related URIs in the exposed data to improve discovery of other related information on the Web.

2 eXtensible Markup Language

1. Well-formedness of the XML

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

Brak zamknięcia tagu 'a'(/)

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

```
a
|-b
| |'foo'
a
| 'bar'
```

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

Brak zamknięcia tagu 'b'()

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

Zła kolejność zamykania tagów.

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

Brakuje zamknięcia głównego tagu na końcu stringu

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

```
a
|-b
| |-c
| | |'foo'
| | |'baz'
| |-c
| |'baz'
```



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

```
a x=1,X=2  
|-b  
| |'foo'
```

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

```
a x=1  
|-b x=2  
| |'foo'
```

```
<a x="1"y="2"x="3"><b>foo</b></a>  
Atrybut x jest dwukrotnie definiowany.
```

```
a x=1,y=2,x=3  
|-b  
| |'foo'
```

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

```
a  
|-b x=3  
| |-c y=1  
| | |'foo'  
| |-c  
| |'bar'
```

2. Creating XML document

```
<book isInteresting="yes">  
  <author>  
    <name>Juan</name>  
    <surname>Kowolsky</surname>  
    <nationality>American</nationality>  
  </author>  
  <title>  
    <originalTitle>House</originalTitle>  
    <polishTitle>Dom</polishTitle>  
    <translator>  
      <name>Jan</name>  
      <surname>Kowalski</surname>  
      <nationality>Poles</nationality>  
    </translator>  
  </title>  
  <date>2001-10-26T09:00:00</date>  
  <publisher>The MIT Press</publisher>  
  <price>31</price>  
</book>
```

3. DTD

```
<?xml version="1.0" ?>  
<!DOCTYPE book [  
<!ELEMENT author ( name | nationality | surname )* >  
<!ELEMENT book ( author, title, date, publisher, price ) >  
<!ELEMENT title ( originalTitle, polishTitle, translator ) >  
<!ELEMENT translator ( name, surname, nationality ) >  
<!ATTLIST book isInteresting CDATA #REQUIRED>  
<!ELEMENT date ( #PCDATA ) >
```

```

<!ELEMENT name ( #PCDATA ) >
<!ELEMENT nationality ( #PCDATA ) >
<!ELEMENT originalTitle ( #PCDATA ) >
<!ELEMENT polishTitle ( #PCDATA ) >
<!ELEMENT price ( #PCDATA ) >
<!ELEMENT publisher ( #PCDATA ) >
<!ELEMENT surname ( #PCDATA ) >
]>

```

4. XSD

```

<?xml version="1.0" encoding="utf-16"?>
<xsd:schema attributeFormDefault="unqualified" elementFormDefault="qualified" version="1.0" xmlns:xsd="http://www.w3.org/2001/XMLSchema" >
  <xsd:element name="book">
    <xsd:complexType>
      <xsd:sequence>
        <xsd:element name="author">
          <xsd:complexType>
            <xsd:sequence>
              <xsd:element name="name" type="xsd:string" />
              <xsd:element name="surname" type="xsd:string" />
              <xsd:element name="nationality" type="xsd:string" />
            </xsd:sequence>
          </xsd:complexType>
        </xsd:element>
        <xsd:element name="title">
          <xsd:complexType>
            <xsd:sequence>
              <xsd:element name="originalTitle" type="xsd:string" />
              <xsd:element name="polishTitle" type="xsd:string" />
              <xsd:element name="translator">
                <xsd:complexType>
                  <xsd:sequence>
                    <xsd:element name="name" type="xsd:string" />
                    <xsd:element name="surname" type="xsd:string" />
                    <xsd:element name="nationality" type="xsd:string" />
                  </xsd:sequence>
                </xsd:complexType>
              </xsd:element>
            </xsd:sequence>
          </xsd:complexType>
        </xsd:element>
        <xsd:element name="date" type="xsd:dateTime" />
        <xsd:element name="publisher" type="xsd:string" />
        <xsd:element name="price" type="xsd:int" />
      </xsd:sequence>
      <xsd:attribute name="isInteresting" type="xsd:string" />
    </xsd:complexType>
  </xsd:element>
</xsd:schema>

```