



Informatyka Stosowana rok IV

Semantic Web

Sprawozdanie 1- 25.04.2012

Łukasz Jaromi
Łukasz Krzyżek

Grupa: Środa 8:00

8 maja 2012

Extracting RDF metadata – Annotating images

Piłkarze Wisły Kraków



Triples of the Data Model

| Nr | Subject | Predicate | Object |
|----|---|---|---|
| 1 | http://www.w3.org/RDF/Validator/run/1336229671705 | http://www.w3.org/2000/10/annotation-ns#annotates | http://cdn17.se.smcloud.net/t/pics/thumbnails/2010/11/05/EN_00939761_0010_640x360_crop_rozmiar-niestandardowy.jpg |
| 2 | http://www.w3.org/RDF/Validator/run/1336229671705 | http://www.w3.org/2000/10/annotation-ns#created | "2012-05-05T14:45:16Z" |
| 3 | http://cdn17.se.smcloud.net/t/pics/thumbnails/2010/11/05/EN_00939761_0010_640x360_crop_rozmiar-niestandardowy.jpg | http://www.w3.org/1999/02/22-rdf-syntax-ns#type | http://xmlns.com/foaf/0.1/Image |
| 4 | http://cdn17.se.smcloud.net/t/pics/thumbnails/2010/11/05/EN_00939761_0010_640x360_crop_rozmiar-niestandardowy.jpg | http://purl.org/dc/elements/1.1/title | "Piłkarze Wisły Kraków" |
| 5 | http://cdn17.se.smcloud.net/t/pics/thumbnails/2010/11/05/EN_00939761_0010_640x360_crop_rozmiar-niestandardowy.jpg | http://purl.org/dc/elements/1.1/description | "Świętujący po zwycięskim meczu piłkarze Wisły Kraków." |
| 6 | http://cdn17.se.smcloud.net/t/pics/thumbnails/2010/11/05/EN_00939761_0010_640x360_crop_rozmiar-niestandardowy.jpg | http://purl.org/dc/elements/1.1/creator | "Adam Kowalski" |

| | | | |
|----|---|---|---|
| 7 | http://cdn17.se.smcloud.net/t/pics/thumbnails/2010/11/05/EN_00939761_0010_640x360_crop_rozmiar-niestandardowy.jpg | http://purl.org/dc/elements/1.1/date | "20-03-2012" |
| 8 | http://cdn17.se.smcloud.net/t/pics/thumbnails/2010/11/05/EN_00939761_0010_640x360_crop_rozmiar-niestandardowy.jpg | http://jibbering.com/vocabs/image/#width | "640" |
| 9 | http://cdn17.se.smcloud.net/t/pics/thumbnails/2010/11/05/EN_00939761_0010_640x360_crop_rozmiar-niestandardowy.jpg | http://jibbering.com/vocabs/image/#height | "360" |
| 10 | http://www.w3.org/RDF/Validator/run/1336229671705#p1 | http://www.w3.org/1999/02/22-rdf-syntax-ns#type | http://jibbering.com/vocabs/image/#Rectangle |
| 11 | http://cdn17.se.smcloud.net/t/pics/thumbnails/2010/11/05/EN_00939761_0010_640x360_crop_rozmiar-niestandardowy.jpg | http://jibbering.com/vocabs/image/#hasPart | http://www.w3.org/RDF/Validator/run/1336229671705#p1 |
| 12 | http://www.w3.org/RDF/Validator/run/1336229671705#p1 | http://jibbering.com/vocabs/image/#points | "429,43 554,162" |
| 13 | http://www.w3.org/RDF/Validator/run/1336229671705#p1 | http://purl.org/dc/elements/1.1/title | "Małecki" |
| 14 | http://www.w3.org/RDF/Validator/run/1336229671705#p1 | http://purl.org/dc/elements/1.1/description | "Patryk Małecki na głównym planie zdjęcia, na tle stadionu Wisły Kraków." |
| 15 | http://www.w3.org/RDF/Validator/run/1336229671705#p1 | http://jibbering.com/vocabs/image/#depicts | genid:A70666 |
| 16 | genid:A70666 | http://purl.org/dc/elements/1.1/description | "Patryk Małecki jest piłkarzem Wisły Kraków, grającym jako napastnik." |
| 17 | http://www.w3.org/RDF/Validator/run/1336229671705#p2 | http://www.w3.org/1999/02/22-rdf-syntax-ns#type | http://jibbering.com/vocabs/image/#Rectangle |
| 18 | http://cdn17.se.smcloud.net/t/pics/thumbnails/2010/11/05/EN_00939761_0010_640x360_crop_rozmiar-niestandardowy.jpg | http://jibbering.com/vocabs/image/#hasPart | http://www.w3.org/RDF/Validator/run/1336229671705#p2 |
| 19 | http://www.w3.org/RDF/Validator/run/1336229671705#p2 | http://jibbering.com/vocabs/image/#points | "227,20 326,141" |

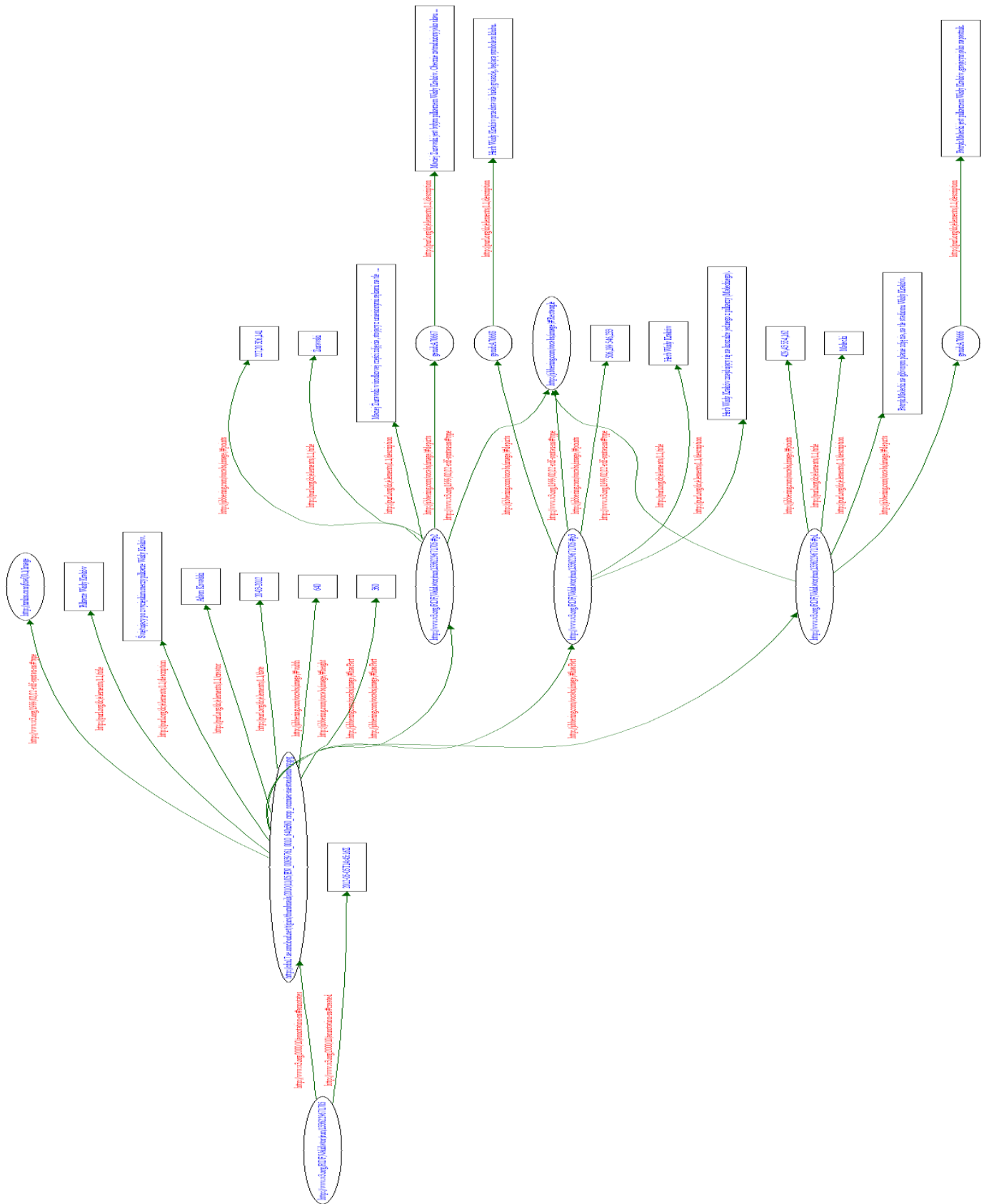
| | | | |
|----|---|---|---|
| | http://www.w3.org/RDF/Validator/run/1336229671705#p2 | http://purl.org/dc/elements/1.1/titl | |
| 20 | http://www.w3.org/RDF/Validator/run/1336229671705#p2 | http://purl.org/dc/elements/1.1/titl | "Żurawski" |
| 21 | http://www.w3.org/RDF/Validator/run/1336229671705#p2 | http://purl.org/dc/elements/1.1/description | "Maciej Żurawski w środkowej części zdjęcia, stojący z uniesionymi rękami na tle stadionu Wisły Kraków." |
| 22 | http://www.w3.org/RDF/Validator/run/1336229671705#p2 | http://jibbering.com/vocabs/image/#depicts | genid:A70667 |
| 23 | genid:A70667 | http://purl.org/dc/elements/1.1/description | "Maciej Żurawski jest byłym piłkarzem Wisły Kraków. Obecnie zatrudniony jako skaut." |
| 24 | http://www.w3.org/RDF/Validator/run/1336229671705#p3 | http://www.w3.org/1999/02/22-rdf-syntax-ns#type | http://jibbering.com/vocabs/image/#Rectangle |
| 25 | http://cdn17.se.smccloud.net/t/pics/thumbnails/2010/11/05/EN_00939761_0010_640x360_crop_rozmiar-niestandardowy.jpg | http://jibbering.com/vocabs/image/#hasPart | http://www.w3.org/RDF/Validator/run/1336229671705#p3 |
| 26 | http://www.w3.org/RDF/Validator/run/1336229671705#p3 | http://jibbering.com/vocabs/image/#points | "506,199 546,233" |
| 27 | http://www.w3.org/RDF/Validator/run/1336229671705#p3 | http://purl.org/dc/elements/1.1/titl | "Herb Wisły Kraków" |
| 28 | http://www.w3.org/RDF/Validator/run/1336229671705#p3 | http://purl.org/dc/elements/1.1/description | "Herb Wisły Kraków znajdujący się na koszulce jednego z piłkarzy (Małeckiego)." |
| 29 | http://www.w3.org/RDF/Validator/run/1336229671705#p3 | http://jibbering.com/vocabs/image/#depicts | genid:A70668 |
| 30 | genid:A70668 | http://purl.org/dc/elements/1.1/description | "Herb Wisły Kraków przedstawia białą gwiazdę, będącą symbolem klubu." |

The original RDF/XML document

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: <rdf:Description rdf:about="">
10:   <an:annotates
rdf:resource="http://cdn17.se.smcloud.net/t/pics/thumbnails/2010/11/05/EN_00939761_0010_640x360_crop_rozmiar-niestandardowy.jpg"/>
11:   <an:created>2012-05-05T14:45:16Z</an:created>
12: </rdf:Description>
13: <foaf:Image
rdf:about="http://cdn17.se.smcloud.net/t/pics/thumbnails/2010/11/05/EN_00939761_0010_640x360_crop_rozmiar-niestandardowy.jpg">
14:   <dc:title>Piłkarze Wisły Kraków</dc:title>
15:   <dc:description>Świętujący po zwycięskim meczy piłkarze Wisły Kraków.</dc:description>
16:   <dc:creator>Adam Kowalski</dc:creator>
17:   <dc:date>20-03-2012</dc:date>
18:   <image:width>640</image:width>
19:   <image:height>360</image:height>
20:   <image:hasPart>
21:     <image:Rectangle rdf:ID="p1">
22:       <image:points>429,43 554,162</image:points>
23:       <dc:title>Małecki</dc:title>
24:       <dc:description>Patryk Małecki na głównym planie zdjęcia, na tle stadionu Wisły Kraków.
</dc:description>
25:       <image:depicts rdf:parseType="Resource">
26:         <dc:description>Patryk Małecki jest piłkarzem Wisły Kraków, grającym jako napastnik.
</dc:description>
27:       </image:depicts>
28:     </image:Rectangle>
29:   </image:hasPart>
30:   <image:hasPart>
31:     <image:Rectangle rdf:ID="p2">
32:       <image:points>227,20 326,141</image:points>
33:       <dc:title>Żurawski</dc:title>
34:       <dc:description>Maciej Żurawski w środkowej części zdjęcia, stojący z uniesionymi rękami na tle
stadionu Wisły Kraków. </dc:description>
35:       <image:depicts rdf:parseType="Resource">
36:         <dc:description>Maciej Żurawski jest byłym piłkarzem Wisły Kraków. Obecnie zatrudniony jako skaut.
</dc:description>
37:       </image:depicts>
38:     </image:Rectangle>
39:   </image:hasPart>
40:   <image:hasPart>
41:     <image:Rectangle rdf:ID="p3">
42:       <image:points>506,199 546,233</image:points>
43:       <dc:title>Herb Wisły Kraków</dc:title>
44:       <dc:description>Herb Wisły Kraków znajdujący się na koszulce jednego z piłkarzy
(Małeckiego).</dc:description>
45:       <image:depicts rdf:parseType="Resource">
46:         <dc:description>Herb Wisły Kraków przedstawia białą gwiazdę, będącą symbolem klubu.
</dc:description>
47:       </image:depicts>
48:     </image:Rectangle>
49:   </image:hasPart>
50: </foaf:Image>
51: </rdf:RDF>
```

Graph of the data model



Common Vocabularies

SKOS - (Simple Knowledge Organization System) - jest rodziną słowników służących do reprezentowania podstawowych struktur, treści (zbiorów semantycznych i hierarchicznie powiązanych terminów) typu: tezaursus, classification scheme, taksonomii (nauka o zasadach i metodach klasyfikowania oraz tworzenia i opisywani jednostek systematycznych) lub procedur kontroli autorytatywnej.

Wykorzystywany przez: *Library of Congress Subject Headings, GeoNames.*

Dublin Core - ogólny standard metadanych rozwijany przez Dublin Core Metadata Initiative. Służy do opisu zasobów internetowych. DC definiuje 15 prostych elementów (Title, Creator, Subject, Description, Publisher, Contributor, Date, Type, Format, Identifier, Source, Language, Relation, Coverage, Rights).

Wykorzystywany przez: *Open Source Metadata Framework, PBCore, aplikacje Zope CMF's.*

FOAF - jeden z ustandaryzowanych schematów formatowania ze zdefiniowanym przeznaczeniem. FOAF służy do opisu danej osoby. FOAF, podobnie jak RDF, charakteryzuje się tym, że może być przetwarzany przez maszyny. Dodatkowo dzięki możliwości odwołania się do innych dokumentów FOAF (w praktyce do opisów innych osób), tworzy się sieć społeczną (ang. social network). Specyfikacja FOAF opiera się na składni XML i strukturze RDF.

Wykorzystywany przez: *aplikacje, serwisy internetowe, działające w obszarze social media (np. Social Graph API).*

DOAP - służy do opisu projektów programistycznych/informatycznych. Udostępniony na licencji open-source.

Wykorzystywany przez: *Mozilla Foundation.*

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 Krzyżek</foaf:name>
<foaf:title>Mr</foaf:title>
<foaf:givenname>Łukasz</foaf:givenname>
<foaf:family_name>Krzyżek</foaf:family_name>
<foaf:nick>lukkrz</foaf:nick>
<foaf:mbox_sha1sum>7db680b5e578c90a86bcd1c62a8b0185f3653980</foaf:mbox_sha1sum>
<foaf:phone rdf:resource="tel:124000000"/>
<foaf:workplaceHomepage rdf:resource="www.efestiwal.net"/>
<foaf:schoolHomepage rdf:resource="http://www.viii-lo.krakow.pl"/></foaf:Person>
</rdf:RDF>
```

Link: <http://www.efestiwal.net/foaf/foaf.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: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 Jaromi</foaf:name>
```

```
<foaf:title>Mr</foaf:title>
<foaf:givenname>Łukasz</foaf:givenname>
<foaf:family_name>Jaromi</foaf:family_name>
<foaf:mbox_sha1sum>a847107a05600bca7f25fb6d7d7aabaddc13ac85</foaf:mbox_sha1sum>
<foaf:phone rdf:resource="tel:120000000"/></foaf:Person>
</rdf:RDF>
```

Link: <http://www.efestiwal.net/foaf/foaf2.rdf>

Schemas

Ontologie powstały w latach 2000 – 2003, ich rozmiar jest bardzo zróżnicowany. Złożoność i zagnieżdżenie (klas i podklas) tworzonych ontologii zależy przede wszystkim od problemu (rzeczy) jakie chcemy opisać.

Ontology visualization

Aplikacja "OwlSight" jest niedostępna.

Ontology reasoning

Aplikacja "OwlSight" jest niedostępna.

Linked Open Data

LinkedGeoData używa informacji zebrane przez OpenStreetMap Project i udostępnia je w formacie RDF (w ramach LinkedData).

DBpedia zawiera ustrukturyzowane dane z Wikipedii i umożliwia zadawanie zapytań korzystając z Semantic Web.

flickr wrappr rozszerzenie DBpedia zawierające linki do zdjęć zamieszczonych w flickr.

Calais narzędzie umożliwiające dodanie semantycznego zapisu danych do blogów, CMSów lub innych stron.

LinkedMDB zawiera dane związane z filmografią.

Semantic Web tools

RDFizers -> narzędzie do automatycznego generowania reprezentacji danych w formacie RDF

Użyteczne formaty, które powinien również obsługiwać RDFizers to przede wszystkim inne formaty graficzne (np. png).

Semantic Web browsers -> aplikacje umożliwiające przeglądanie Linked Data

Obecnie funkcjonuje 13 przeglądarek.

Semantic Web Search Engines -> wyszukiwarki semantic web

Pomimo faktu, że spora część z wymienionych wyszukiwarek już nie istnieje (not found) to korzystanie z działających (np. Falcons i swoogle) jest bardzo łatwe i szybkie.

Applications

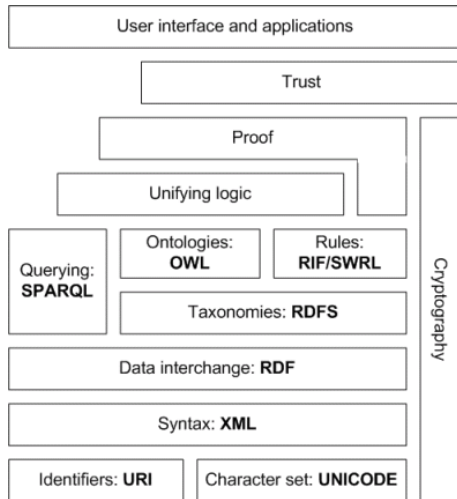
Bardzo przydatnym narzędziem jest RDFizers. Naszym zdaniem największe perspektywy rozwoju daje DBpedia.

Control questions

1) Można dodać adnotacje korzystając z gotowych aplikacji (np. RDFizers) lub dodać ją do naszej strony ręcznie.

2) Semantic Web Stack ilustruje architekturę (budowę, stos) Semantic Web.

Warstwy wchodzące w skład architektury możemy podzielić na:



3) Wyrażenia w RDF złożone są z:

- temat (podmiot) - znacznik RDF URI lub pusty węzeł
- predykat - znacznik RDF URI
- obiekt - znacznik RDF URI, literał lub pusty węzeł

Główną zaletą składni RDF jest jej czytelność i prostota w użyciu.

4) Ontologie w Semantic Web pełnią rolę strukturalnego schematu do organizowania informacji (danych) w postaci obiektów i relacji między nimi.

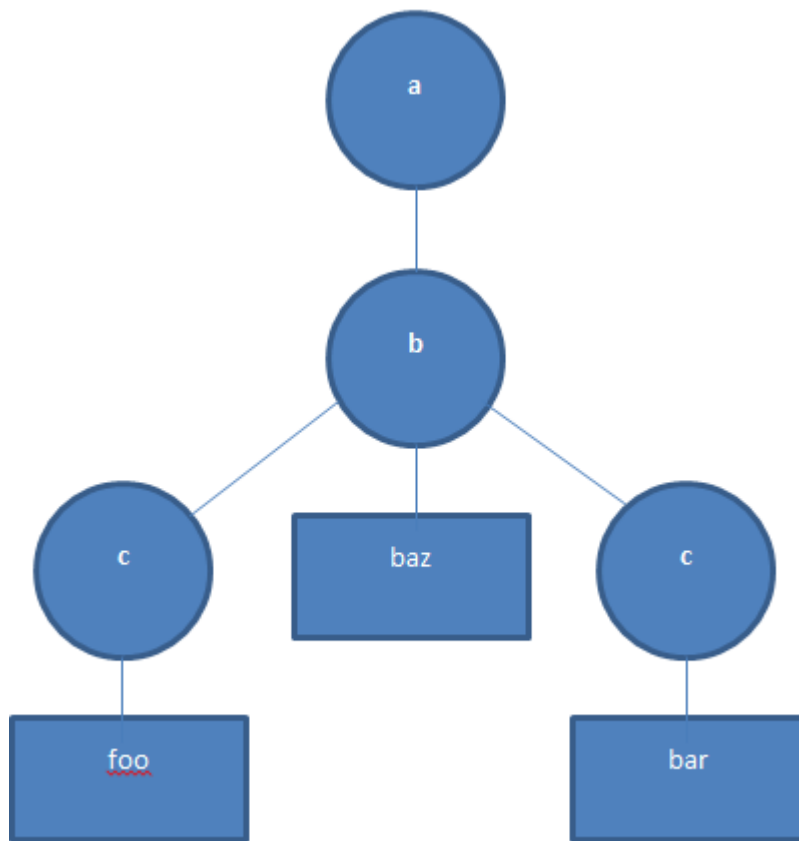
5) 4 zasady Linked Data. Należy:

- używać URI do identyfikowania (nazwy) dla obiektów
- używać HTTP URI aby uzyskać informacje o tych obiektach (przez ludzi i agentów)
- aktywacja poszczególnego URI ma spowodować pozyskanie użytecznych informacji (standardowy format RDF/XML)
- wykorzystuj powiązania z innymi obiektami za pomocą URI, aby zapewnić możliwość eksploracji i odkrywania innych obiektów w sieci.

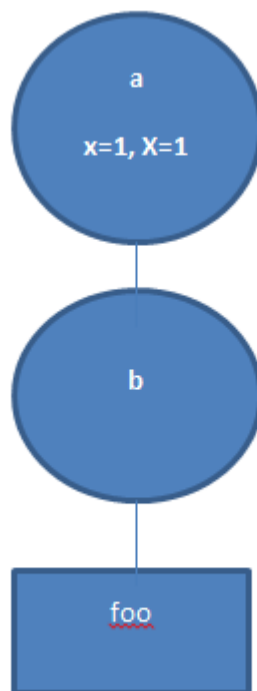
Warm-up: Well-formedness of the XML

- `<a><a>` -> Błąd. Znacznik `<a>` nie został zamknięty.
- `<a>foo<a>bar` Błąd. Brak nadrzędnego elementu (znacznika root).
- `<a>foobar` -> Błąd. Drugi znacznik `` nie został zamknięty.
- `<a>foobar` -> Błąd. Kolejność otwierania i zamykania znaczników musi być zachowana.
- `<a/>foobar` -> Błąd. Drugi znacznik `` nie został zamknięty.

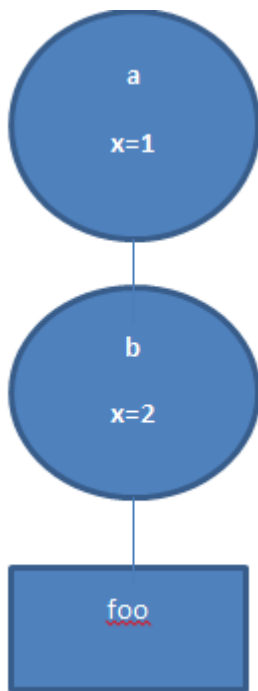
- <a><c>foo</c>baz<c>bar</c> -> ok.



- foo -> ok.

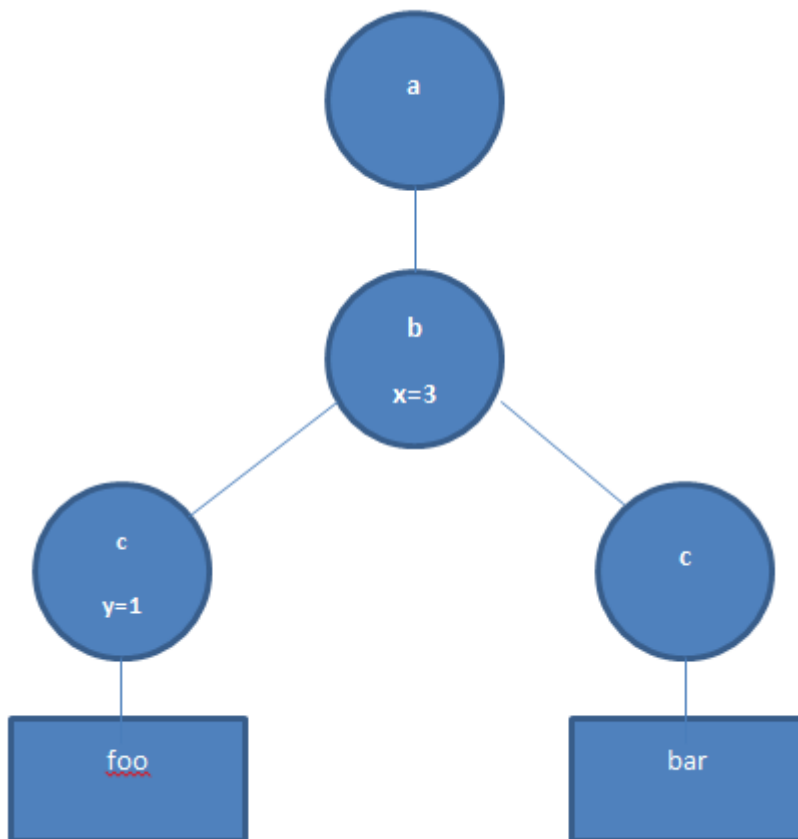


- `<b x="2">foo` -> ok.



- `foo` -> Błąd. Ten sam atrybyt 'x' użyty dwa razy.

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



Creating XML document

```
<bookshelf>
  <book>
    <author>Witold Gombrowicz</author>
    <title>Ferdydurke</title>
    <date>2008-11-09</date>
    <publisher>PWN</publisher>
    <favourite>yes</favourite>
  </book>
</bookshelf>
```

DTD

```
<?xml version="1.0"?>
<!DOCTYPE bookshelf [
<!ELEMENT book (author,title,date,publisher,favourite)>
<!ELEMENT author (#PCDATA)>
<!ELEMENT title (#PCDATA)>
<!ELEMENT date (#PCDATA)>
<!ELEMENT publisher (#PCDATA)>
<!ELEMENT favourite (#PCDATA)>
]>
<bookshelf>
  <book>
    <author>Witold Gombrowicz</author>
    <title>Ferdydurke</title>
    <date>2008-11-09</date>
    <publisher>PWN</publisher>
    <favourite>yes</favourite>
  </book>
</bookshelf>
```

XSD

```
<?xml version="1.0"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema">

<xs:element name="bookshelf">
  <xs:complexType>
    <xs:sequence>
      <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="favourite" type="xs:string" default="no"/>
          </xs:sequence>
        </xs:complexType>
      </xs:element>
    </xs:sequence>
  </xs:complexType>
</xs:element>

</xs:schema>
```