

Akademia Górniczo-Hutnicza
Wydział EAIiE
Informatyka Stosowana, I Rok II Stopnia

**Sprawozdanie 1:
Introduction to the Semantic Web,
eXtensible Markup Language (XML)**

**Krzysztof Kutt
Michał Nowak**

**Podstawy Inżynierii Wiedzy
Środa 14:00**

Kraków 2012

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#"
>
  <rdf:Description rdf:about="">
    <an:annotates rdf:resource="http://www.kanzaki.com/works/2004/
imgdsc/miniduck_by_ccc.jpg"/>
    <an:author>Krzysiek i Michal</an:author>
    <an:created>2012-05-08T21:09:26Z</an:created>
  </rdf:Description>
  <foaf:Image rdf:about="http://www.kanzaki.com/works/2004/imgdsc/
miniduck_by_ccc.jpg">
    <dc:title>Trzy kaczki</dc:title>
    <dc:description>Trzy zolte sztuczne kaczuszki</dc:description>
    <image:width>300</image:width>
    <image:height>225</image:height>
    <image:hasPart>
      <image:Rectangle rdf:ID="p1">
        <image:points>151,93 205,132</image:points>
        <dc:title>Dziob</dc:title>
        <dc:description>Dziob kaczki</dc:description>
      </image:Rectangle>
    </image:hasPart>
    <image:hasPart>
      <image:Rectangle rdf:ID="p2">
        <image:points>214,95 233,114</image:points>
        <dc:title>Oko</dc:title>
        <dc:description>Oko kaczki</dc:description>
      </image:Rectangle>
    </image:hasPart>
    <image:hasPart>
      <image:Rectangle rdf:ID="p3">
        <image:points>249,90 282,200</image:points>
        <dc:title>Kuper</dc:title>
        <dc:description>Kuper kaczki</dc:description>
      </image:Rectangle>
    </image:hasPart>
    <image:hasPart>
      <image:Rectangle rdf:ID="p4">
        <image:points>23,30 139,162</image:points>
        <dc:title>Kaczka 2</dc:title>
        <dc:description>Druga kaczka</dc:description>
      </image:Rectangle>
    </image:hasPart>
  </foaf:Image>
</rdf:RDF>
```

```

<image:hasPart>
  <image:Rectangle rdf:ID="p5">
    <image:points>141,10 220,74</image:points>
    <dc:title>Kaczka 3</dc:title>
    <dc:description>Trzecia kaczka</dc:description>
  </image:Rectangle>
</image:hasPart>
</foaf:Image>
</rdf:RDF>

```

Nr	Subject	Predicate	Object
1	http://www.w3.org/RDF/Validator/run/1336512462763	http://www.w3.org/2000/10/annotation-ns#annotates	http://www.kanzaki.com/works/2004/imgdsc/miniduck_by_ccc.jpg
2	http://www.w3.org/RDF/Validator/run/1336512462763	http://www.w3.org/2000/10/annotation-ns#author	"Krzysiek i Michal"
3	http://www.w3.org/RDF/Validator/run/1336512462763	http://www.w3.org/2000/10/annotation-ns#created	"2012-05-08T21:09:26Z"
4	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
5	http://www.kanzaki.com/works/2004/imgdsc/miniduck_by_ccc.jpg	http://purl.org/dc/elements/1.1/title	"Trzy kaczki"
6	http://www.kanzaki.com/works/2004/imgdsc/miniduck_by_ccc.jpg	http://purl.org/dc/elements/1.1/description	"Trzy zolte sztuczne kaczuszki"
7	http://www.kanzaki.com/works/2004/imgdsc/miniduck_by_ccc.jpg	http://jibbering.com/vocabs/image/#width	"300"
8	http://www.kanzaki.com/works/2004/imgdsc/miniduck_by_ccc.jpg	http://jibbering.com/vocabs/image/#height	"225"
9	http://www.w3.org/RDF/Validator/run/1336512462763#p1	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/1336512462763#p1
11	http://www.w3.org/RDF/Validator/run/1336512462763#p1	http://jibbering.com/vocabs/image/#points	"151,93 205,132"

12	http://www.w3.org/RDF/Validator/run/1336512462763#p1	http://purl.org/dc/elements/1.1/title	"Dziob"
13	http://www.w3.org/RDF/Validator/run/1336512462763#p1	http://purl.org/dc/elements/1.1/description	"Dziob kaczuki"
14	http://www.w3.org/RDF/Validator/run/1336512462763#p2	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://jibbering.com/vocabs/image/#Rectangle
15	http://www.kanzaki.com/works/2004/imgdsc/miniduck_by_ccc.jpg	http://jibbering.com/vocabs/image/#hasPart	http://www.w3.org/RDF/Validator/run/1336512462763#p2
16	http://www.w3.org/RDF/Validator/run/1336512462763#p2	http://jibbering.com/vocabs/image/#points	"214,95 233,114"
17	http://www.w3.org/RDF/Validator/run/1336512462763#p2	http://purl.org/dc/elements/1.1/title	"Oko"
18	http://www.w3.org/RDF/Validator/run/1336512462763#p2	http://purl.org/dc/elements/1.1/description	"Oko kaczuki"
19	http://www.w3.org/RDF/Validator/run/1336512462763#p3	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://jibbering.com/vocabs/image/#Rectangle
20	http://www.kanzaki.com/works/2004/imgdsc/miniduck_by_ccc.jpg	http://jibbering.com/vocabs/image/#hasPart	http://www.w3.org/RDF/Validator/run/1336512462763#p3
21	http://www.w3.org/RDF/Validator/run/1336512462763#p3	http://jibbering.com/vocabs/image/#points	"249,90 282,200"
22	http://www.w3.org/RDF/Validator/run/1336512462763#p3	http://purl.org/dc/elements/1.1/title	"Kuper"
23	http://www.w3.org/RDF/Validator/run/1336512462763#p3	http://purl.org/dc/elements/1.1/description	"Kuper kaczuki"
24	http://www.w3.org/RDF/Validator/run/1336512462763#p4	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/1336512462763#p4

26	http://www.w3.org/RDF/Validator/run/1336512462763#p4	http://jibbering.com/vocabs/image/#points	"23,30 139,162"
27	http://www.w3.org/RDF/Validator/run/1336512462763#p4	http://purl.org/dc/elements/1.1/title	"Kaczka 2"
28	http://www.w3.org/RDF/Validator/run/1336512462763#p4	http://purl.org/dc/elements/1.1/description	"Druga kaczka"
29	http://www.w3.org/RDF/Validator/run/1336512462763#p5	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://jibbering.com/vocabs/image/#Rectangle
30	http://www.kanzaki.com/works/2004/imgdsc/miniduck_by_ccc.jpg	http://jibbering.com/vocabs/image/#hasPart	http://www.w3.org/RDF/Validator/run/1336512462763#p5
31	http://www.w3.org/RDF/Validator/run/1336512462763#p5	http://jibbering.com/vocabs/image/#points	"141,10 220,74"
32	http://www.w3.org/RDF/Validator/run/1336512462763#p5	http://purl.org/dc/elements/1.1/title	"Kaczka 3"
33	http://www.w3.org/RDF/Validator/run/1336512462763#p5	http://purl.org/dc/elements/1.1/description	"Trzecia kaczka"

2 Common Vocabularies

- **SKOS**. *Simple Knowledge Organization System*. Rodzina języków formalnych przeznaczonych do reprezentacji tezaurusów, systemów klasyfikacji, taksonomii lub innego rodzaju słowników o kontrolowanej strukturze.
Wykorzystywany m.in. w: AGROVOC, GEMET, Library of Congress Subject Headings, GeoNames.
- **Dublin Core**. Ogólny standard metadanych, składający się z 15 prostych elementów (tytuł, twórca, opis, słowa kluczowe, etc). Przyjęty jako standard ISO.
Wykorzystywany m.in. w: Open Source Metadata Framework, dLibra.
- **FOAF**. *Friend of a Friend*. Format służący do opisu osób. Dzięki możliwości odwoływania się do innych dokumentów FOAF można tworzyć sieci społeczne.
Wykorzystywany m.in. w: Google Social Graph API, Yahoo SearchMonkey, Yandex
- **DOAP**. *Description of a Project*. Format służący do opisu projektów informatycznych.
Wykorzystywany m.in. w: projektach Mozilla Foundation, Freshmeat.

3. FOAF

Michał [<http://student.agh.edu.pl/~kkutt/piw2012/michal.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>Michał Nowak</foaf:name>
<foaf:title>Mr</foaf:title>
<foaf:givenname>Michał</foaf:givenname>
<foaf:family_name>Nowak</foaf:family_name>
<foaf:nick>Nowaksson</foaf:nick>
<foaf:mbox_shalsum>078df3c13932045e337e1bd799752f548e1a2c56</foaf:mbox_shalsum>
<foaf:schoolHomepage rdf:resource="www.agh.edu.pl"/>
<foaf:knows>
<foaf:Person>
<foaf:name>Krzysztof</foaf:name>
<foaf:mbox_shalsum>ef6961ce298ff0fe70564d3c733c802c5d0114df</foaf:mbox_shalsum></foaf:Person></foaf:knows></foaf:Person>
</rdf:RDF>
```

Krzysiek [<http://student.agh.edu.pl/~kkutt/piw2012/krzysiek.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>Krzysztof Kutt</foaf:name>
<foaf:title>Mr</foaf:title>
<foaf:givenname>Krzysztof</foaf:givenname>
<foaf:family_name>Kutt</foaf:family_name>
<foaf:nick>Rubik</foaf:nick>
<foaf:mbox_sha1sum>668484c9c2cd169f71ad4215135c93f2b6b2f2a9</foaf:mbox_sha1sum>
<foaf:homepage rdf:resource="http://student.agh.edu.pl/~kkutt/">
<foaf:schoolHomepage rdf:resource="http://www.agh.edu.pl/">
<foaf:knows>
<foaf:Person>
<foaf:name>Michał Nowak</foaf:name>
<foaf:mbox_sha1sum>fb2f0e79ebb21f75b0cd42085c051061d7078dfc</foaf:mbox_sha1sum>
<rdfs:seeAlso rdf:resource="http://student.agh.edu.pl/~kkutt/piw2012/michal.rdf"/></foaf:Person></foaf:knows></foaf:Person>
</rdf:RDF>
```

4 Schemas

Ontologie różnią się rokiem powstania i wielkością. Im później ontologia powstała tym więcej informacji zawiera.

5 Ontology visualisation

Nie udało nam się uruchomić aplikacji.

6 Ontology reasoning

Nie udało nam się uruchomić aplikacji.

7 Linked Open Data

- UK Postcodes [<http://www.uk-postcodes.com/>]. Informacje o kodach pocztowych w Wielkiej Brytanii
- Poképédia [<http://www.pokepedia.fr/>]. Zbiór informacji o Pokemonach
- IEEE [<http://ieee.rkbexplorer.com/>]. Informacje o artykułach z bazy IEEE
- GeoNames [<http://www.geonames.org/>]. Informacje o nazwach geograficznych / topograficznych (alternatywne nazwy, wysokość n.p.m., współrzędne, etc)
- BBC Music [<http://www.bbc.co.uk/music>]. Duża baza informacji o muzyce (płyty, utwory, wydarzenia, etc)

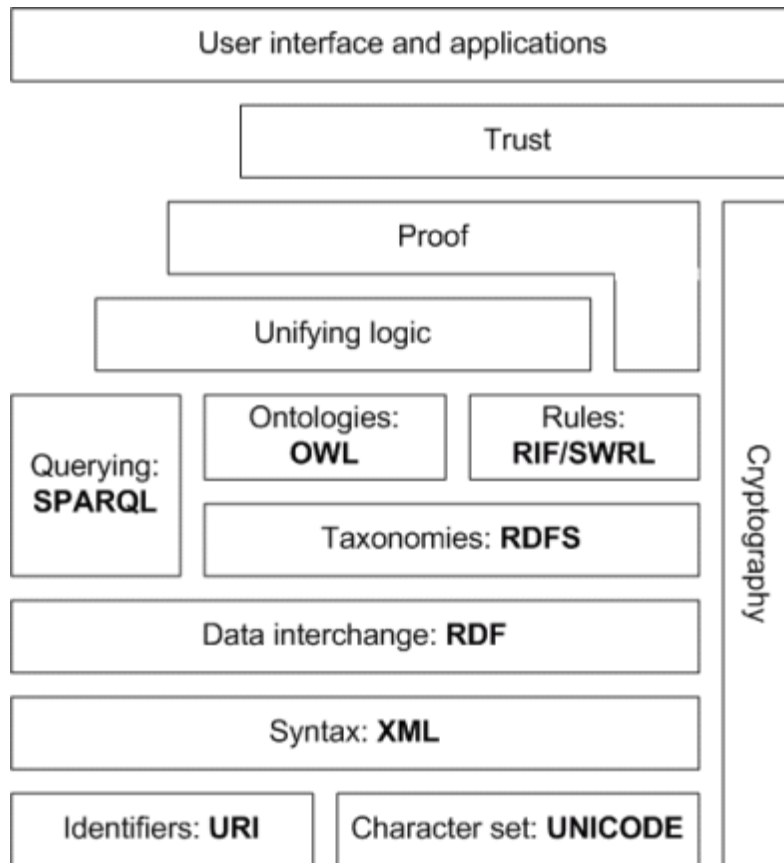
8 Semantic Web tools

- **RDFizers: which in your opinion are the most important? What an RDFizer would you add (converting from what format)?**
Najważniejsze wg nas konwertery to te związane z kalendarzem. Sprowadzenie wszystkich kalendarzy do jednego formatu RDF znacząco pomaga zapanować nad wszystkimi zaplanowanymi wydarzeniami.
Dobrze by było dodać konwertery dla plików multimedialnych (np mp3), które pomogły by zebrać informacje o plikach niezależnie od ich typów.
- **Semantic Web browsers: How many of them are up and running?**
 - działają [9]: Tabulator, Disco, OpenLink Data Explorer, Zitgist RDF Browser, DBpedia Mobile, Marbles, zLinks, Sigma, Quick & Dirty RDF Browser
 - nie działają [4]: razorbase, iLOD, Fenfire (tylko PDF z krótką informacją o projekcie), Objectviewer
- **Semantic Web Search Engines: How easy is it to get started with each of them?**
Bardzo proste. Wystarczy w odpowiednim oknie wpisać nazwę szukanego elementu (lub jego URI) i kliknąć Search. Silnik po chwili zwraca informacje powiązane z tym słowem.
- **Applications of the Linked Data, including mashups and everyday tools: Which ones do you find useful? What other applications can you think of?**
Spodobała nam się aplikacja LinkedGeoData Browser, która jest wielką mapą świata z zaznaczonymi punktami z informacjami w formacie RDF o danym punkcie. Dużo podróżujemy i dzięki tej aplikacji możemy łatwo zbierać informacje o odwiedzanych miejscach.

Control questions

- **How one can add semantic annotation to a web page?**
Na dwa główne sposoby: albo ręcznie wpisać adnotacje w kodzie, albo skorzystać z odpowiednich generatorów (i przekleić adnotacje do kodu).

- Explain the Semantic Web Stack of technologies.**
 Stos Semantic Web przedstawia hierarchię zależności pomiędzy technologiami / językami. Każda kolejna warstwa zależy od poprzedniej.
 Stos pokazuje w jaki sposób Sieci Semantyczne są zorganizowane, aby mogły działać wg założeń.

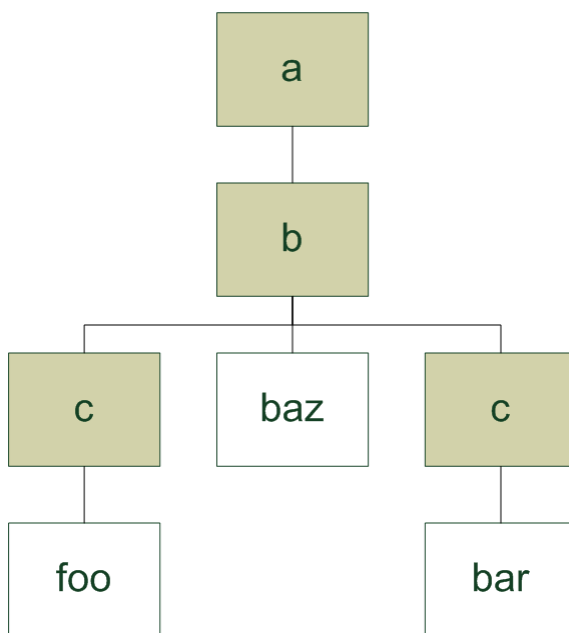


- What is the main syntax for RDF? What are its advantages over other syntaxes?**
 RDF jest oparty na formacie XML, który jest wspierany przez większość współczesnych języków programowania i do którego powstało wiele narzędzi.
- What is the role of the ontologies in the Semantic Web?**
 Ontologie wyszczególniają elementy dziedziny i relacje między nimi. Służą jako podstawa wnioskowania.
- What are the 4 principles of Linked Data?**
 - Używanie URI do identyfikowania rzeczy
 - Możliwość wyszukiwania i odnoszenia się do danej rzeczy, przez ludzi i agentów, dzięki dostępowi do URI przez HTTP
 - Dostarczanie użytecznych informacji o danej rzeczy, przy wykorzystaniu standardowych formatów, takich jak RDF/XML
 - Załączanie odnośników do powiązanych URI, aby ułatwić wyszukiwanie informacji w sieci

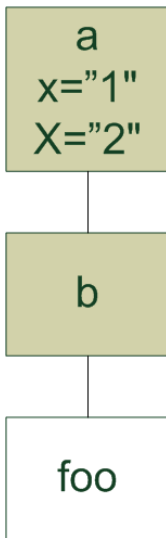
eXtensible Markup Language (XML)

1 Warm-up: Well-formedness of the XML

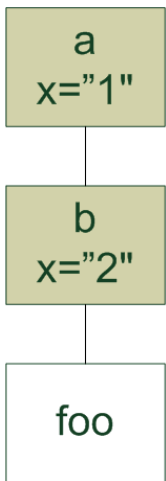
- `<a><a>`
Błędny XML. Brak tagów zamykających dla `<a>`
- `<a>foo<a>bar`
Błędny XML. Może istnieć tylko jeden element nadrzędny
- `<a>foobar`
Błędny XML. Brak tagu zamykającego dla ``
- `<a><c>foobar</c>`
Błędny XML. Niewłaściwa kolejność zamykania tagów `` i `<c>`
- `<a/>foobar`
Błędny XML. Brak tagów zamykających dla ``. Może istnieć tylko jeden element nadrzędny.
- `<a><c>foo</c>baz<c>bar</c>`
Poprawny XML.



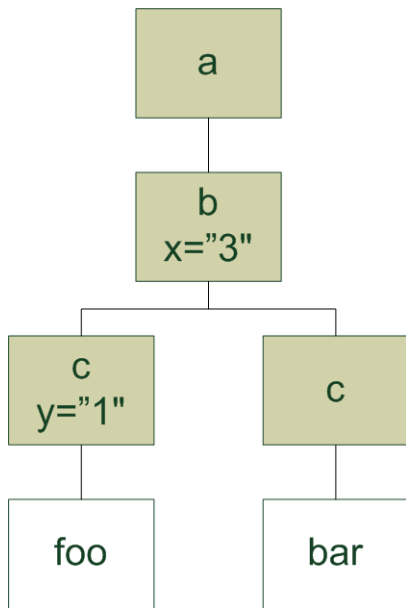
- `foo`
Poprawny XML.



- `<b x="2">foo`
Poprawny XML.



- `foo`
Błędny XML. Powtórna deklaracja atrybutu x w tagu <a>
- `<a><b x="3"><c y="1">foo</c><c>bar</c>`
Poprawny XML.



2 Creating XML document

```

<ksiazka>
  <autor>Stieg Larrson</autor>
  <tytul>Mężczyźni, którzy nienawidzą kobiet</tytul>
  <seria>Millenium</seria>
  <tlumacz>Beata Walczak-Larsson</tlumacz>
  <wydawnictwo>Czarna Owca</wydawnictwo>
  <rok>2009</rok>
  <miejsce>Warszawa</miejsce>
</ksiazka>

```

3 DTD

```

<?xml version="1.0" ?>
<!DOCTYPE ksiazka [
<!ELEMENT ksiazka
(autor*,tytul,seria?,tlumacz*,wydawnictwo,rok,miejsce)>
  <!ELEMENT autor (#PCDATA)>
  <!ELEMENT tytul (#PCDATA)>
  <!ELEMENT seria (#PCDATA)>
  <!ELEMENT tlumacz (#PCDATA)>
  <!ELEMENT wydawnictwo (#PCDATA)>
  <!ELEMENT rok (#PCDATA)>
  <!ELEMENT miejsce (#PCDATA)>
]>

```

5 XSD

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

<xs:element name="ksiazka">
  <xs:complexType>
    <xs:sequence>
      <xs:element name="autor" type="xs:string" minOccurs="1"
maxOccurs="unbounded"/>
      <xs:element name="tytul" type="xs:string"/>
      <xs:element name="seria" type="xs:string" minOccurs="0"
maxOccurs="1"/>
      <xs:element name="tlumacz" type="xs:string" minOccurs="0"
maxOccurs="unbounded"/>
      <xs:element name="wydawnictwo" type="xs:string"/>
      <xs:element name="rok" type="xs:gYear"/>
      <xs:element name="miejsce" type="xs:string"/>
    </xs:sequence>
  </xs:complexType>
</xs:element>

</xs:schema>
```