

# Podstawy Inżynierii Wiedzy

Informatyka Stosowana

Wydział Elektrotechniki, Automatyki,  
Informatyki i Elektroniki



# AGH

## Semantic Web

*Sprawozdanie 1- 25.04.2012*

Marek Walczyk  
Monika Krupnik

Grupa: Środa 8:00

## Introduction to the Semantic Web

2.

Image:



The resulting RDF:

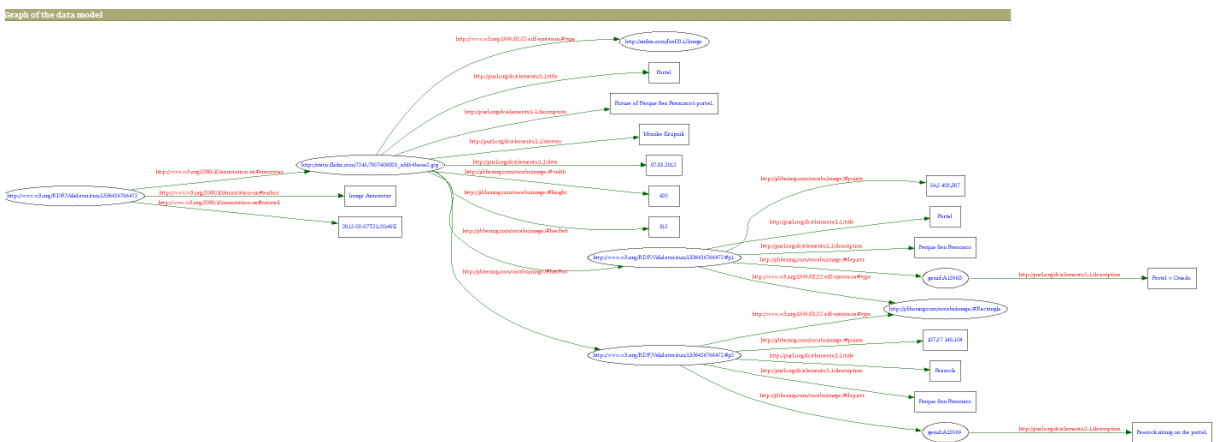
```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:foaf="http://xmlns.com/foaf/0.1/"
  xmlns:image="http://jibbering.com/vocabs/image/#"
  xmlns:an="http://www.w3.org/2000/10/annotation-ns#"
>
  <rdf:Description rdf:about="">
    <an:annotates rdf:resource="http://static.flickr.com/7245/7007400820_a3d94beaa2.jpg"/>
    <an:author>Image Annotator</an:author>
    <an:created>2012-05-07T21:20:49Z</an:created>
  </rdf:Description>
  <foaf:Image rdf:about="http://static.flickr.com/7245/7007400820_a3d94beaa2.jpg">
    <dc:title>Portal</dc:title>
    <dc:description>Picture of Parque San Francisco's portal.</dc:description>
    <dc:creator>Monika Krupnik</dc:creator>
    <dc:date>07.05.2012</dc:date>
    <image:width>420</image:width>
    <image:height>315</image:height>
    <image:hasPart>
      <image:Rectangle rdf:ID="p1">
```

```

<image:points>54,5 409,307</image:points>
<dc:title>Portal</dc:title>
<dc:description>Parque San Francisco</dc:description>
<image:depicts rdf:parseType="Resource">
  <dc:description>Portal w Oviedo</dc:description>
</image:depicts>
</image:Rectangle>
</image:hasPart>
<image:hasPart>
  <image:Rectangle rdf:ID="p2">
    <image:points>107,57 140,104</image:points>
    <dc:title>Peacock</dc:title>
    <dc:description>Parque San Francisco</dc:description>
    <image:depicts rdf:parseType="Resource">
      <dc:description>Peacock sitting on the portal.</dc:description>
    </image:depicts>
  </image:Rectangle>
</image:hasPart>
</foaf:Image>
</rdf:RDF>

```

Triples of the Data Model			
Number	Subject	Predicate	Object
1	<a href="http://www.w3.org/RDF/Validator/run/1336426766472">http://www.w3.org/RDF/Validator/run/1336426766472</a>	<a href="http://www.w3.org/2000/10/annotation-ns#annotates">http://www.w3.org/2000/10/annotation-ns#annotates</a>	<a href="http://static.flickr.com/7245/7007400820_a3d94beaa2.jpg">http://static.flickr.com/7245/7007400820_a3d94beaa2.jpg</a>
2	<a href="http://www.w3.org/RDF/Validator/run/1336426766472">http://www.w3.org/RDF/Validator/run/1336426766472</a>	<a href="http://www.w3.org/2000/10/annotation-ns#author">http://www.w3.org/2000/10/annotation-ns#author</a>	"Image Annotator"
3	<a href="http://www.w3.org/RDF/Validator/run/1336426766472">http://www.w3.org/RDF/Validator/run/1336426766472</a>	<a href="http://www.w3.org/2000/10/annotation-ns#created">http://www.w3.org/2000/10/annotation-ns#created</a>	"2012-05-07T21:20:49Z"
4	<a href="http://static.flickr.com/7245/7007400820_a3d94beaa2.jpg">http://static.flickr.com/7245/7007400820_a3d94beaa2.jpg</a>	<a href="http://www.w3.org/1999/02/22-rdf-syntax-ns#type">http://www.w3.org/1999/02/22-rdf-syntax-ns#type</a>	<a href="http://xmlns.com/foaf/0.1/Image">http://xmlns.com/foaf/0.1/Image</a>
5	<a href="http://static.flickr.com/7245/7007400820_a3d94beaa2.jpg">http://static.flickr.com/7245/7007400820_a3d94beaa2.jpg</a>	<a href="http://purl.org/dc/elements/1.1/title">http://purl.org/dc/elements/1.1/title</a>	"Portal"
6	<a href="http://static.flickr.com/7245/7007400820_a3d94beaa2.jpg">http://static.flickr.com/7245/7007400820_a3d94beaa2.jpg</a>	<a href="http://purl.org/dc/elements/1.1/description">http://purl.org/dc/elements/1.1/description</a>	"Picture of Parque San Francisco's portal."
7	<a href="http://static.flickr.com/7245/7007400820_a3d94beaa2.jpg">http://static.flickr.com/7245/7007400820_a3d94beaa2.jpg</a>	<a href="http://purl.org/dc/elements/1.1/creator">http://purl.org/dc/elements/1.1/creator</a>	"Monika Krupnik"
8	<a href="http://static.flickr.com/7245/7007400820_a3d94beaa2.jpg">http://static.flickr.com/7245/7007400820_a3d94beaa2.jpg</a>	<a href="http://purl.org/dc/elements/1.1/date">http://purl.org/dc/elements/1.1/date</a>	"07.05.2012"
9	<a href="http://static.flickr.com/7245/7007400820_a3d94beaa2.jpg">http://static.flickr.com/7245/7007400820_a3d94beaa2.jpg</a>	<a href="http://jibbering.com/vocabs/image/#width">http://jibbering.com/vocabs/image/#width</a>	"420"
10	<a href="http://static.flickr.com/7245/7007400820_a3d94beaa2.jpg">http://static.flickr.com/7245/7007400820_a3d94beaa2.jpg</a>	<a href="http://jibbering.com/vocabs/image/#height">http://jibbering.com/vocabs/image/#height</a>	"315"
11	<a href="http://www.w3.org/RDF/Validator/run/1336426766472#p1">http://www.w3.org/RDF/Validator/run/1336426766472#p1</a>	<a href="http://www.w3.org/1999/02/22-rdf-syntax-ns#type">http://www.w3.org/1999/02/22-rdf-syntax-ns#type</a>	<a href="http://jibbering.com/vocabs/image/#Rectangle">http://jibbering.com/vocabs/image/#Rectangle</a>
12	<a href="http://static.flickr.com/7245/7007400820_a3d94beaa2.jpg">http://static.flickr.com/7245/7007400820_a3d94beaa2.jpg</a>	<a href="http://jibbering.com/vocabs/image/#hasPart">http://jibbering.com/vocabs/image/#hasPart</a>	<a href="http://www.w3.org/RDF/Validator/run/1336426766472#p1">http://www.w3.org/RDF/Validator/run/1336426766472#p1</a>
13	<a href="http://www.w3.org/RDF/Validator/run/1336426766472#p1">http://www.w3.org/RDF/Validator/run/1336426766472#p1</a>	<a href="http://purl.org/dc/elements/1.1/title">http://purl.org/dc/elements/1.1/title</a>	"54,5 409,307"
14	<a href="http://www.w3.org/RDF/Validator/run/1336426766472#p1">http://www.w3.org/RDF/Validator/run/1336426766472#p1</a>	<a href="http://purl.org/dc/elements/1.1/description">http://purl.org/dc/elements/1.1/description</a>	"Parque San Francisco"
15	<a href="http://www.w3.org/RDF/Validator/run/1336426766472#p1">http://www.w3.org/RDF/Validator/run/1336426766472#p1</a>	<a href="http://jibbering.com/vocabs/image/#depicts">http://jibbering.com/vocabs/image/#depicts</a>	genid:A15968
16	<a href="http://www.w3.org/RDF/Validator/run/1336426766472#p1">http://www.w3.org/RDF/Validator/run/1336426766472#p1</a>	<a href="http://purl.org/dc/elements/1.1/description">http://purl.org/dc/elements/1.1/description</a>	"Portal w Oviedo"
17	genid:A15968	<a href="http://jibbering.com/vocabs/image/#Rectangle">http://jibbering.com/vocabs/image/#Rectangle</a>	<a href="http://www.w3.org/RDF/Validator/run/1336426766472#p2">http://www.w3.org/RDF/Validator/run/1336426766472#p2</a>
18	<a href="http://www.w3.org/RDF/Validator/run/1336426766472#p2">http://www.w3.org/RDF/Validator/run/1336426766472#p2</a>	<a href="http://www.w3.org/1999/02/22-rdf-syntax-ns#type">http://www.w3.org/1999/02/22-rdf-syntax-ns#type</a>	<a href="http://jibbering.com/vocabs/image/#Rectangle">http://jibbering.com/vocabs/image/#Rectangle</a>
19	<a href="http://static.flickr.com/7245/7007400820_a3d94beaa2.jpg">http://static.flickr.com/7245/7007400820_a3d94beaa2.jpg</a>	<a href="http://jibbering.com/vocabs/image/#hasPart">http://jibbering.com/vocabs/image/#hasPart</a>	<a href="http://www.w3.org/RDF/Validator/run/1336426766472#p2">http://www.w3.org/RDF/Validator/run/1336426766472#p2</a>
20	<a href="http://www.w3.org/RDF/Validator/run/1336426766472#p2">http://www.w3.org/RDF/Validator/run/1336426766472#p2</a>	<a href="http://jibbering.com/vocabs/image/#points">http://jibbering.com/vocabs/image/#points</a>	"107,57 140,104"
21	<a href="http://www.w3.org/RDF/Validator/run/1336426766472#p2">http://www.w3.org/RDF/Validator/run/1336426766472#p2</a>	<a href="http://purl.org/dc/elements/1.1/title">http://purl.org/dc/elements/1.1/title</a>	"Peacock"
22	<a href="http://www.w3.org/RDF/Validator/run/1336426766472#p2">http://www.w3.org/RDF/Validator/run/1336426766472#p2</a>	<a href="http://purl.org/dc/elements/1.1/description">http://purl.org/dc/elements/1.1/description</a>	"Parque San Francisco"
23	<a href="http://www.w3.org/RDF/Validator/run/1336426766472#p2">http://www.w3.org/RDF/Validator/run/1336426766472#p2</a>	<a href="http://jibbering.com/vocabs/image/#depicts">http://jibbering.com/vocabs/image/#depicts</a>	genid:A15969
24	genid:A15969	<a href="http://purl.org/dc/elements/1.1/description">http://purl.org/dc/elements/1.1/description</a>	"Peacock sitting on the portal."



## 2) Common Vocabularies

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

Programy wykorzystujące SKOS:


- Tezaursus w SWED Environmental Directory.
- ontologia GeoName do reprezentacji danych geograficznych.

**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, pośrednio Gnome i KDE.
- PBCore – rozszerza DC o elementy do opisu contentu audio-wideo.
- Zope Content Management Framework – framework do zarządzania contentem, pozwala na katalogowanie, personalizację contentu.

**FOAF (Friend of a friend)** - 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. Wykorzystanie FOAF:

- Platformy blogowe: Live Journal, DeadJournal wspierają powiązania pomiędzy
- użytkownikami wykorzystując FOAF
- Społeczność MyOpera, związana z przeglądarka Opera
-  Yandex – wyszukiwanie blogów

**DOAP (Description of a Project)** – schemat RDF wraz ze znacznikami xml, służące do opisu projektów informatycznych (głównie Open Source). Wykorzystywany w projektach Mozilla Foundation oraz na stronie Freecode (dawniej Freshmeat), umożliwiającej wyszukiwanie oprogramowania.

### 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>Krzysiu Krawczyk</foaf:name>
<foaf:title>Major</foaf:title>
<foaf:givenname>Krzysiu</foaf:givenname>
<foaf:family_name>Krawczyk</foaf:family_name>
<foaf:nick>Czacha</foaf:nick>
<foaf:mbox_sha1sum>b01b5835fa8ae7b7582968a7ecacb9b85503a6c9</foaf:mbox_sha1sum>
<foaf:homepage rdf:resource="www.homepage.com"/>
<foaf:depiction rdf:resource="brak"/>
<foaf:phone rdf:resource="tel:+48125946358"/>
<foaf:workplaceHomepage rdf:resource="Dom"/>
<foaf:workInfoHomepage rdf:resource="nic"/>
<foaf:schoolHomepage rdf:resource="AGH"/>
<foaf:knows>
<foaf:Person>
<foaf:name>Mietek</foaf:name>
<foaf:mbox_sha1sum>fa65a6d11c9117d5359a3534004c278a9b931929</foaf:mbox_sha1sum></foaf:Person></foaf:knows></foaf:Person>
</rdf:RDF>
```

Link do pliku rdf: <http://www.shr3ku.pl/download/foaf.rdf>

Wizualizacja:

# Krzysiu Krawczyk

## Knows:

Mietek  *sha1sum of a personal mailbox URI name:*  
fa65a6d11c9117d5359a3534004c278a9b931929

**schoolHomepage:**

<http://www.shr3ku.pl/download/AGH>

**work info homepage:**

<http://www.shr3ku.pl/download/nic>

**workplace homepage:**

<http://www.shr3ku.pl/download/Dom>

**phone:**

<tel:+48125946358>

**depiction:**

<http://www.shr3ku.pl/download/brak>

**homepage:**

<http://www.shr3ku.pl/download/www.homepage.com>

**sha1sum of a personal mailbox URI name:**

b01b5835fa8ae7b7582968a7ecacb9b85503a6c9

**nickname:**

Czacha

**family\_name:**

Krawczyk

**Given name:**

Krzysiu

**title:**

Major


### Referenced Vocabularies


-  <http://www.w3.org/1999/02/22-rdf-syntax-ns#>  
The RDF Vocabulary (RDF)
-  <http://xmlns.com/foaf/0.1/>
-  <http://webns.net/mvcb/>

Wizualizacja z dodanym friendem:

## Krzysiu Krawczyk

### Knows:

Czesio  *sha1sum of a personal mailbox URI name:*  
bd0e0916b39bbba064f73abd79c720e6597f5597

Mietek  *sha1sum of a personal mailbox URI name:*  
fa65a6d11c9117d5359a3534004c278a9b931929

#### *schoolHomepage:*

<http://www.shr3ku.pl/download/AGH>

#### *work info homepage:*

<http://www.shr3ku.pl/download/nic>

#### *workplace homepage:*

<http://www.shr3ku.pl/download/Dom>

#### *phone:*

<tel:+48125946358>

#### *depiction:*

<http://www.shr3ku.pl/download/brak>

#### *homepage:*

<http://www.shr3ku.pl/download/www.homepage.com>

#### *sha1sum of a personal mailbox URI name:*

b01b5835fa8ae7b7582968a7ecacb9b85503a6c9

#### *nickname:*

Czacha

#### *family\_name:*

Krawczyk


#### *Given name:*

Krzysiu


#### *title:*


Major

#### Referenced Vocabularies

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

The RDF Vocabulary (RDF)

 <http://xmlns.com/foaf/0.1/>

 <http://webns.net/mvcb/>

4. Ontologie z katalogu powstały w latach 2000 - 2003. Ich wielkość jest różna, większość z nich zawiera kilkanaście - kilkadziesiąt rekordów.

5, 6. Strona programu owl:sight nie jest dostępna podczas tworzenia sprawozdania, co uniemożliwia wykonanie ćwiczeń.

## 7) Linked Open Data

- LinkedGeoData używa informacje zebrane przez OpenStreetMap Project i udostępnia je w formacie RDF (w ramach LinkedData).
- Open Library – otwarty katalog książek, udostępniający użytkownikom możliwości dodawania, edycji informacji o książkach. Mottem twórców „biblioteki” jest: „One web page for every book ever published”. Obecnie w katalogu zgromadzonych jest ok. 20 mln rekordów.
- Mortality EnAkTing – zawiera statystyki dotyczące śmiertelności w Wielkiej Brytanii w latach 2008-2009. Dane pochodzą z UK Home Office (Ministerstwo Spraw Wewnętrznych Wielkiej Brytanii)
- LinkedMDB zawiera dane związane z filmografią.
- Flickr wrappr rozszerzenie DBpedia zawierające linki do zdjęć zamieszczonych w flickr

## 8) Semantic Web tools

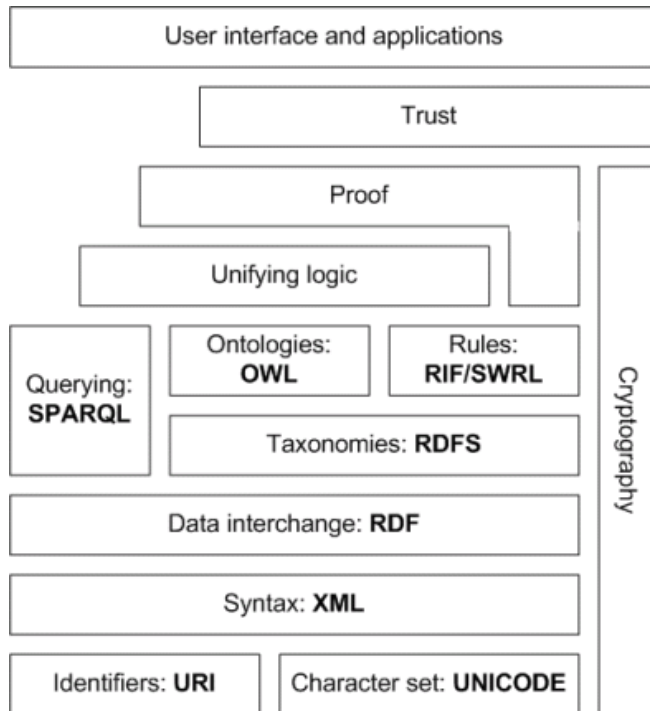
- **RDFizers** - Narzędzie do automatycznego generowania reprezentacji danych w formacie RDF. Użyteczne formaty, które powinien również obsługiwać RDFizers to przede wszystkim inne formaty graficzne (np. png).
- **Semantic Web browsers** - Aplikacje umożliwiające przeglądanie Linked Data. Obecnie funkcjonuje 13 przeglądarek, z których 6 nie działa (Razorbase, Tabulator, Zitgist RDF Browser, Objectviewer, iLOD, zLinks) a 5 działa (Disco, OpenLinkDataExplorer, Dbpedia Mobile, Sigma i Quick & Dirty RDF Browser).
- **Semantic Web Search Engines** - W większości działających mechanizmów wyszukiwujących wystarczy wpisać nazwę firmy/przedmiotu/osobę, o których informacji poszukujemy. Wyszukiwanie przy pomocy „semantycznych” wyszukiwarek od strony użytkownika nie różni się niczym od obecnie wykorzystywanych, standardowych wyszukiwarek internetowych typu Google.
- **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 widać na poniższym rysunku



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) Cztery zasady Linked Data. Należy:

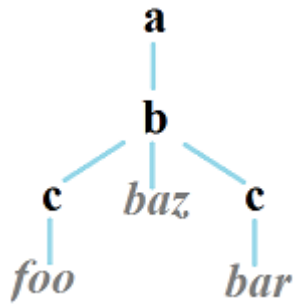
- używać URI do identyfikowania (nazwy) dla obiektów
- używać HTTP URI aby uzyskać informacje o tych obiektach (przez ludzi i agentów)
- aktywacja poszczególnego URI ma spowodować pozyskanie użytecznych informacji (standardowy format RDF/XML)
- wykorzystuj powiązania z innymi obiektami za pomocą URI,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

# XML

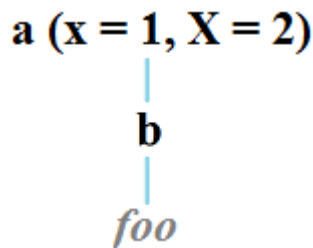
## 1) Warm-up: Well-formedness of the XML

### Poprawne:

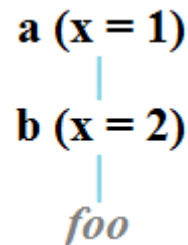
- `<a><b><c>foo</c>baz<c>bar</c></b></a>`



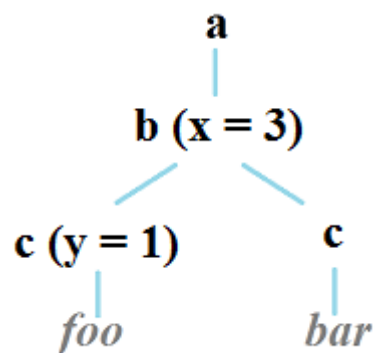
- `<a x="1" X="2"><b>foo</b></a>`



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



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



### Niepoprawne:

- `<a><b>foo</b></a><a>bar</a>` - brak korzenia dokumentu

- `<a><b>foo</b><b>bar</a>` - niezamknięty element `<b>`
- `<a><b><c>foo</b>bar</c></a>` - niepoprawna kolejność tagów otwierających i zamykających
- `<a/><b>foo</b><b>bar<b>` - brak korzenia dokumentu
- `<a x="1" y="2" x="3"><b>foo</b></a>` - podwójna deklaracja atrybutu x

## Creating XML document

```

    <?xml version="1.0" ?>
<!DOCTYPE book [
<!ELEMENT book (author,title,date,publisher)>
<!ELEMENT author (#PCDATA)>
<!ELEMENT title (#PCDATA)>
<!ELEMENT date (#PCDATA)>
<!ELEMENT publisher (#PCDATA)>
]>
<book>
  <author>Tom Clancy</author>
  <title>Kolekcjoner</title>
  <date>2004</date>
  <publisher>Wydawnictwo De Facto</publisher>
</book>

```

## XSD

```

<?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:positiveInteger"/>
        <xs:element name="publisher" type="xs:string"/>
      </xs:sequence>
    </xs:complexType>
  </xs:element>

</xs:schema>

```