

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

Wydział Elektrotechniki, Automatyki, Informatyki i Elektroniki

KATEDRA AUTOMATYKI



SPRAWOZDANIE

TOMASZ DRZEWIECKI, JOANNA GAJDA

SPRAWOZDANIE 1: INTRO, XML

Kraków 2012

Spis treści

1. Intro	3
1.1. Extracting RDF metadata	3
1.2. Common Vocabularies.....	5
1.2.1. SKOS.....	5
1.2.2. Dublin Core.....	5
1.2.3. FOAF.....	6
1.2.4. DOAP.....	6
1.3. FOAF.....	6
1.4. Schemas.....	8
1.5. Ontology visualisation.....	8
1.6. Ontology reasoning	8
1.7. Linked Open Data.....	10
1.8. Semantic Web tools	10
1.8.1. RDFizers	10
1.8.2. Semantic Web browsers	10
1.8.3. Semantic Web Search Engines.....	10
1.8.4. Applications of the Linked Data	10
1.9. Control Questions.....	11
1.9.1. How one can add semantic annotation to a web page?.....	11
1.9.2. Explain the Semantic Web Stack of technologies.....	11
1.9.3. What is the main syntax for RDF? What are its advantages over other syntaxes?	11
1.9.4. What is the role of the ontologies in the Semantic Web?	11
1.9.5. What are the 4 principles of Linked Data?	11
2. XML	12
2.1. Warm-up.....	12
2.2. Creating XML Document.....	16
2.3. DTD.....	16
2.4. XSD.....	16

1. Intro

1.1. Extracting RDF metadata

Wynik walidatora: „Your RDF document validated successfully.”

Trójki:

Number	Subject	Predicate	Object
1	http://www.w3.org/RDF/Validator/run/1336492199211	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/1336492199211	http://www.w3.org/2000/10/annotation-ns#author	“Asia i Tomek”
3	http://www.w3.org/RDF/Validator/run/1336492199211	http://www.w3.org/2000/10/annotation-ns#created	“2012-05-08T15:45:04Z”
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	“Sztuczne kaczki”
6	http://www.kanzaki.com/works/2004/imgdsc/miniduck_by_ccc.jpg	http://purl.org/dc/elements/1.1/description	“Zdjęcie przedstawia trzy sztuczne kaczki.”
7	http://www.kanzaki.com/works/2004/imgdsc/miniduck_by_ccc.jpg	http://purl.org/dc/elements/1.1/creator	“nieznany”
8	http://www.kanzaki.com/works/2004/imgdsc/miniduck_by_ccc.jpg	http://purl.org/dc/elements/1.1/date	“08.05.2012”
9	http://www.kanzaki.com/works/2004/imgdsc/miniduck_by_ccc.jpg	http://jibbering.com/vocabs/image/#width	“300”

10	http://www.kanzaki.com/-works/2004/imgdsc/-miniduck_by_ccc.jpg	http://jibbering.com/vocabs/-image/#height	“225”
11	http://www.w3.org/-RDF/Validator/run/-1336492199211#p1	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://jibbering.com/vocabs/-image/#Rectangle
12	http://www.kanzaki.com/-works/2004/imgdsc/-miniduck_by_ccc.jpg	http://jibbering.com/vocabs/-image/#hasPart	http://www.w3.org/-RDF/Validator/run/-1336492199211#p1
13	http://www.w3.org/-RDF/Validator/run/-1336492199211#p1	http://jibbering.com/vocabs/-image/#points	“44,30 150,110”
14	http://www.w3.org/-RDF/Validator/run/-1336492199211#p1	http://purl.org/dc/elements/1.1/-title	“Głowa kaczki”
15	http://www.w3.org/-RDF/Validator/run/-1336492199211#p1	http://purl.org/dc/elements/1.1/-description	“Głowa sztucznej kaczki.”
16	http://www.w3.org/-RDF/Validator/run/-1336492199211#p1	http://jibbering.com/vocabs/-image/#depicts	genid:A48335
17	genid:A48335	http://purl.org/dc/elements/1.1/-description	“Sztuczne kaczki są używane do kąpieli.”
18	http://www.w3.org/-RDF/Validator/run/-1336492199211#p2	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://jibbering.com/vocabs/-image/#Rectangle
19	http://www.kanzaki.com/-works/2004/imgdsc/-miniduck_by_ccc.jpg	http://jibbering.com/vocabs/-image/#hasPart	http://www.w3.org/-RDF/Validator/run/-1336492199211#p2
20	http://www.w3.org/-RDF/Validator/run/-1336492199211#p2	http://jibbering.com/vocabs/-image/#points	“164,66 284,204”
21	http://www.w3.org/-RDF/Validator/run/-1336492199211#p2	http://purl.org/dc/elements/1.1/-title	“Kaczka”
22	http://www.w3.org/-RDF/Validator/run/-1336492199211#p2	http://purl.org/dc/elements/1.1/-description	“Sztuczna kaczka.”

23	http://www.w3.org/RDF/Validator/run-1336492199211#p2	http://jibbering.com/vocabs/-image/#depicts	genid:A48336
24	genid:A48336	http://purl.org/dc/elements/1.1/-description	“Sztuczne kaczki są dobre do kąpieli.”

Graf: graf przedstawiowy jest na rysunku 1.1.

1.2. Common Vocabularies

1.2.1. SKOS

Zapewnia model dla wyrażania podstawowej struktury i zawartości słownictw kontrolowanych. Pozwala na udokumentowanie, linkowanie i łączenie z innymi danymi tych zbiorów.

Przykłady aplikacji:

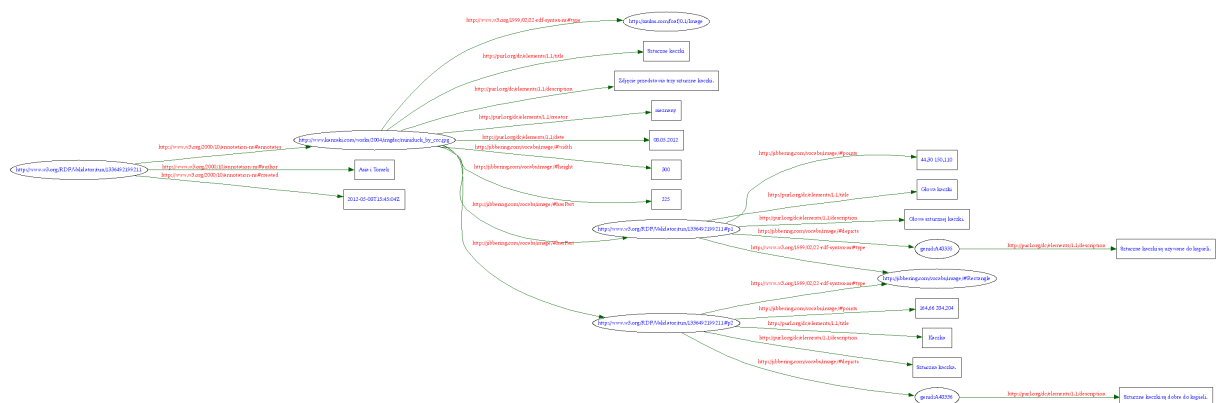
- Lexarus (<http://www.k-int.com/products/lexauruseditor>) - program do zarządzania słownikami również zapisanymi z użyciem SKOS,
- ThManager (<http://thmanager.sourceforge.net>) - program do tworzenia i wizualizacji słowników SKOS.

1.2.2. Dublin Core

Jest to słownik służący do opisu dzieł kulturalnych takich jak książki, albumy muzyczne, filmy.

Przykłady aplikacji:

- Dublin Core Generator (<http://www.dublincoregenerator.com/>) - program do generowania metadanych w standardzie Dublin Core,
- Omeka (<http://omeka.org/>) - program do publikowania archiwum cyfrowego z użyciem standardu Dublin Core.



Rysunek 1.1: Graf dokumentu

1.2.3. FOAF

Projekt poświęcony łączeniu ludzi i informacji poprzez sieć. Służy do opisu osoby, jej podstawowych danych (imię, nazwisko, itp.) oraz gdzie pracuje lub się uczy. Za pomocą FOAF można stworzyć sieć społecznościową.

Przykłady aplikacji:

- Semantic Radar (<https://addons.mozilla.org/en-US/firefox/addon/semantic-radar/>) - służy do powiadamiania użytkownika o tym, że na danej stronie są obecne metadane (m. in. FOAF),
- Drupal (<http://drupal.org/project/foaf>) - dodatek pozwalający na użycie FOAF w programie Drupal.

1.2.4. DOAP

Jest to słownik służący do opisu projektów informatycznych, w szczególności open source.

Przykłady aplikacji:

- GNOME repositories (<http://git.gnome.org/browse/>) - projekt GNOME używa DOAP do opisu projektów w systemie kontroli wersji git,
- Apache Software Foundation (<http://projects.apache.org/>) - używane jest do opisu i zarządzania projektami fundacji.

1.3. FOAF

Tomasz Drzewiecki: <http://student.agh.edu.pl/~tomko/foaf.rdf>

```
<rdf:RDF xmlns:rdf='http://www.w3.org/1999/02/22-rdf-syntax-ns#' xmlns:foaf:
<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>
Tomasz Drzewiecki
</foaf:name>
<foaf:title>
Inz.
</foaf:title>
<foaf:givenname>
```

```

Tomasz
</foaf:givenname>
<foaf:family_name>
Drzewiecki
</foaf:family_name>
<foaf:nick>
tomko
</foaf:nick>
<foaf:mbox_sha1sum>
ea97598497097d47f2090efaf6bce0c480f5978c
</foaf:mbox_sha1sum>
<foaf:schoolHomepage rdf:resource='http://www.agh.edu.pl' />
<foaf:knows>
<foaf:Person>
<foaf:name>
Joanna Gajda
</foaf:name>
<foaf:mbox_sha1sum>
5bf5dea94ea3e9389731b2a2a71e7966f5152417
</foaf:mbox_sha1sum>
<rdfs:seeAlso rdf:resource='http://student.agh.edu.pl/~agajda/foaf.rdf' />
</foaf:Person>
</foaf:knows>
</foaf:Person>
</rdf:RDF>

```

Joanna Gajda: <http://student.agh.edu.pl/~agajda/foaf.rdf>

```

<rdf:RDF xmlns:rdf='http://www.w3.org/1999/02/22-rdf-syntax-ns#' xmlns:r
<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-mat
<admin:errorReportsTo rdf:resource='mailto:leigh@ldodds.com' />
</foaf:PersonalProfileDocument>
<foaf:Person rdf:ID=''me''>
<foaf:name>
Joanna Gajda
</foaf:name>
<foaf:title>
Inz.

```

```
</foaf:title>
<foaf:givenname>
Joanna
</foaf:givenname>
<foaf:family_name>
Gajda
</foaf:family_name>
<foaf:mbox_sha1sum>
fb73faab6cd42a2f1789e4dca6a07eb1ccfa3440
</foaf:mbox_sha1sum>
<foaf:knows>
<foaf:Person>
<foaf:name>
Tomasz Drzewiecki
</foaf:name>
<foaf:mbox_sha1sum>
ea97598497097d47f2090efaf6bce0c480f5978c
</foaf:mbox_sha1sum>
<rdfs:seeAlso rdf:resource=''http://student.agh.edu.pl/~tomko/foaf.rdf''/>
</foaf:Person>
</foaf:knows>
</foaf:Person>
</rdf:RDF>
```

1.4. Schemas

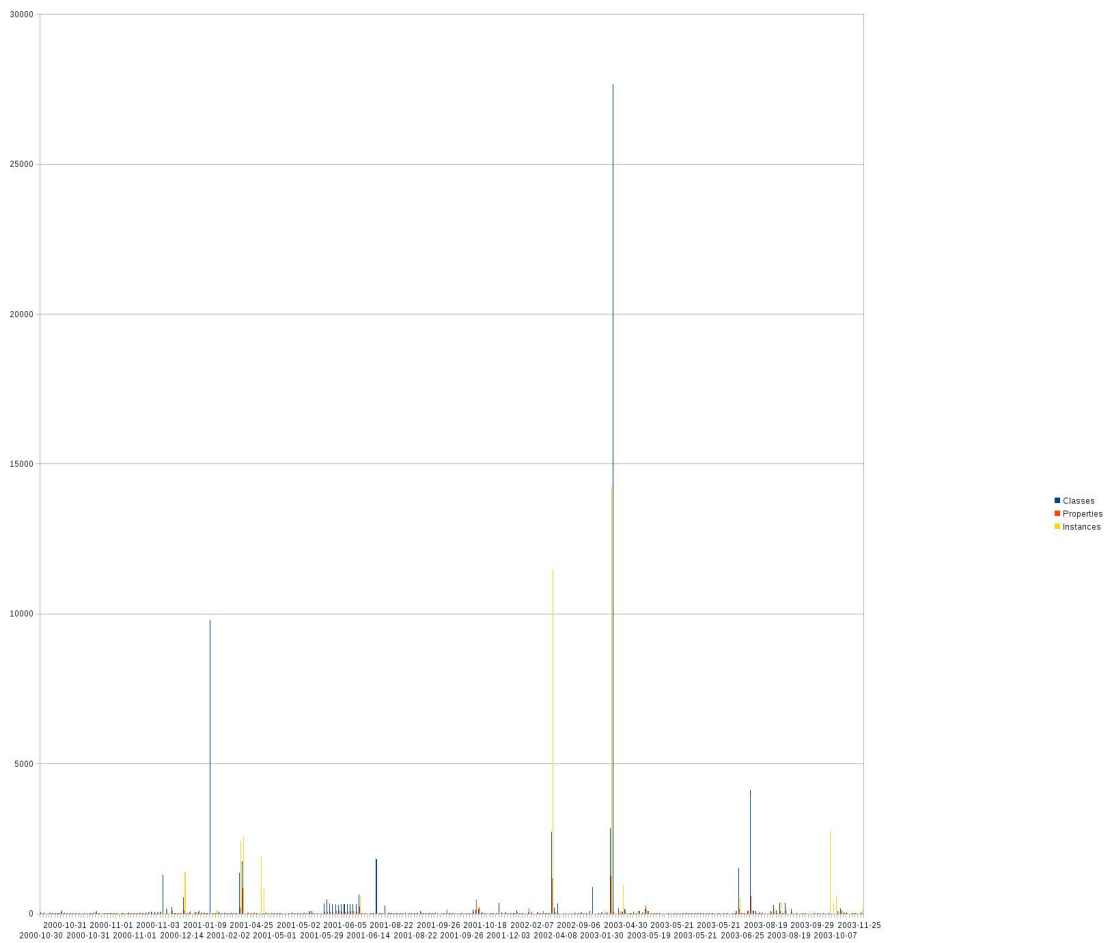
Na wykresie 1.2 przedstawiono rozmiar ontologii w zależności od daty stworzenia. Z wykresu wynika, że rozmiar ontologii rósł i malał w latach 2000-2003. Nie widać żadnego znaczącego trendu w rozmiarze ontologii na przestrzeni czasu.

1.5. Ontology visualisation

Program nie działa.

1.6. Ontology reasoning

Program nie działa.



Rysunek 1.2: Wykres przedstawiający wielkość ontologii

1.7. Linked Open Data

- The Open Library - Otwarta baza książek. Niektóre są dostępne do wypożyczenia.
- BBC Wildlife Finder - Baza informacji o zwierzętach i ich zwyczajach. Dodatkowo też są filmy, zdjęcia, itp.
- ERA - Repozytorium semantyczne zawierające dane linkowane przy pomocy RDF. Dane pochodzą z listy magazynów ARC czytanych przez Australijczyków.
- LIBRIS - Strona pozwalająca na wyszukiwanie książek, muzyki itp. w bazach Szwedzkich bibliotek uniwersyteckich i innych, związanych z badaniami.
- flickr wrappr - Służy do uzupełniania danych umieszczonych w DBpedii linkami rdf do zdjęć umieszczonych na flickr. Dla każdej pozycji w DBpedii generuje kolekcje obrazów z flickr przedstawiające opisywaną pozycję.

1.8. Semantic Web tools

1.8.1. RDFizers

Najistotniejsze mogą być konwertery z Javadoc (zwłaszcza dla programistów), z JPEG do opisu zdjęć oraz z EMail w celu łatwiejszego wyszukiwania istotnych informacji oraz możliwości automatycznego katalogowania.

Warto byłoby stworzyć konwerter z formatów dźwiękowych jak MP3, FLAC. Ułatwiałoby to m. in. wyszukiwanie podobnej muzyki lub wyszukiwania wśród znanych utworów.

1.8.2. Semantic Web browsers

Strona razorbase wygląda na nieczynną. Zitgist RDF Browser nie działa. Po wprowadzeniu URI do własnego pliku FOAF, a także po włączeniu przykładu na stronie głównej występuje błąd. Link do iLOD prowadzi do podglądu katalogu na serwerze. Link do Fenfire prowadzi do pliku pdf z opisem projektu.

Pozostałe przeglądarki wydają się działać. Zwłaszcza te działające on-line jak Marbles i Quick & Dirty RDF Browser, za pomocą których można było obejrzeć własne dokumenty FOAF.

1.8.3. Semantic Web Search Engines

Wszystkie wyszukiwarki oferują prosty w obsłudze interfejs użytkownika, który pozwala na łatwe rozpoczęcie wyszukiwania. Wyszukiwanie podstawowych informacji w niektórych z nich jest proste. Część z wymienionych wyszukiwarek nie działa.

1.8.4. Applications of the Linked Data

Z oferowanych aplikacji na dużą uwagę zasługują BBC Programmes i BBC Music Beta. Ponadto godnym uwagi jest również ODSMail.

Przykładową aplikacją tego rodzaju mógłby być program do zarządzania dokumentami.

1.9. Control Questions

1.9.1. How one can add semantic annotation to a web page?

W celu dodania informacji semantycznych do strony sieci Internet należy w nagłówku dokumentu html umieścić odnośnik do treści semantycznych np. w formacie RDF.

1.9.2. Explain the Semantic Web Stack of technologies.

Na dole stosu znajdują się pliki w formacie Unicode lub URI - identyfikator zasobu. Wyżej znajdują się informacje o powiązaniach pomiędzy konkretnymi zasobami w formacie RDF. Ponad tym znajdują się ontologie określające klasy zasobów oraz powiązania pomiędzy nimi. Wyżej istnieje logika, pozwalająca na wnioskowanie pewnych faktów. Powyżej są dowody. Na samym szczycie znajduje się warstwa zaufania określająca poziom zaufania do poszczególnych źródeł informacji.

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

Najpopularniejszą składnią RDF jest obecnie XML. Jest to spowodowane popularnością XML-a w innych dziedzinach informatyki oraz związaną z tym dużą liczbą bibliotek, programów do przetwarzania XML-a.

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

Rolą ontologii w sieciach semantycznych jest określenie klas obiektów oraz powiązań pomiędzy nimi.

1.9.5. What are the 4 principles of Linked Data?

1. Użycie URI jako identyfikatora obiektu,
2. Użycie protokołu HTTP do określania URI dzięki czemu dane mogą być dostępne fragmentami, dla ludzi i agentów,
3. Dostarczanie użytecznych informacji o obiekcie w przypadku, gdy URI odnosi się do fragmentu obiektu,
4. Dołączanie linków do innych obiektów w celu ulepszenia możliwości odkrywania powiązanych informacji.

2. XML

2.1. Warm-up

– `<a><a>`

Niepoprawne, brak znaczników zamykających dla tagu a.

– `<a>foo <a>bar `

Niepoprawne, dokument XML może zawierać tylko jeden korzeń.

– `<a>foo bar `

Niepoprawne, brak znacznika zamykającego dla tagu b.

– `<a><c>foo bar </c>`

Niepoprawne, znacznik zamykający b występuje przed znacznikiem zamykającym c.

– `<a/>foo bar `

Niepoprawne, dokument XML może zawierać tylko jeden korzeń.

– `<a><c>foo </c>baz<c>bar </c>`

Poprawne drzewo przedstawione na rysunku 2.1.

– `foo `

Poprawne drzewo przedstawione na rysunku 2.2.

– `<b x='2'>foo `

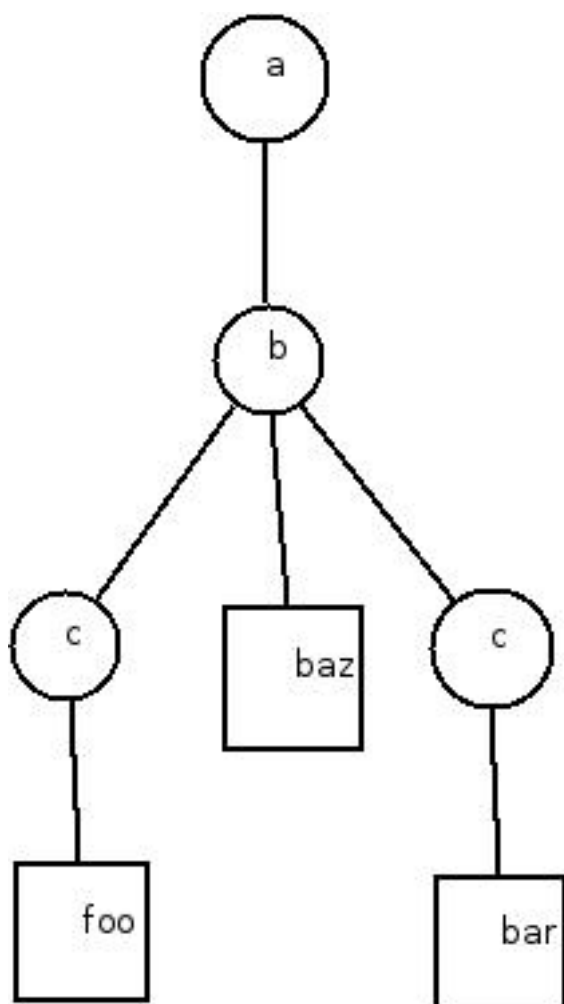
Poprawne drzewo przedstawione na rysunku 2.3.

– `foo `

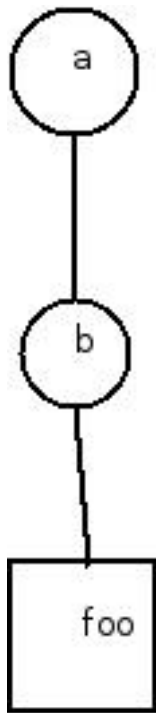
Niepoprawne, atrybut x w znaczniku a występuje dwukrotnie.

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

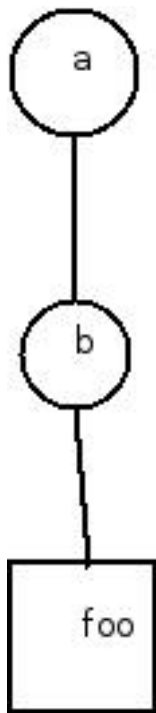
Poprawne drzewo przedstawione na rysunku 2.4.



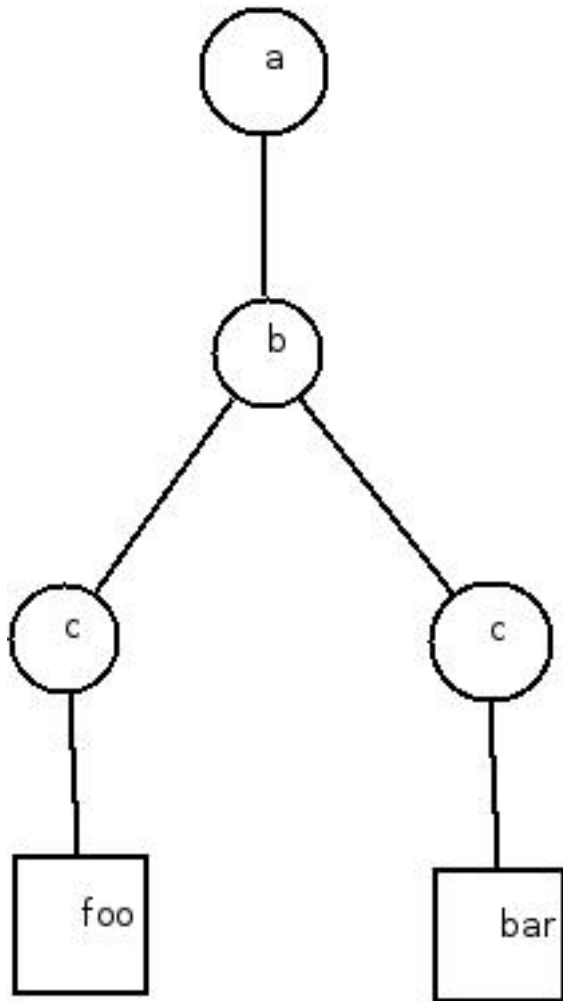
Rysunek 2.1: Drzewo XML



Rysunek 2.2: Drzewo XML



Rysunek 2.3: Drzewo XML



Rysunek 2.4: Drzewo XML

2.2. Creating XML Document

```
<book>
  <author>Terry Pratchett </author>
  <title>Blask Fantastyczny </title>
  <date>1995</date>
  <translator>Piotr W. Cholewa</translator>
  <publisher>Proszynski i S-ka</publisher>
</book>
```

2.3. DTD

```
<!DOCTYPE book [

<!ELEMENT book (author, title, date, translator, publisher)>
<!ELEMENT author (#PCDATA)>
<!ELEMENT title (#PCDATA)>
<!ELEMENT date (#PCDATA)>
<!ELEMENT translator (#PCDATA)>
<!ELEMENT publisher (#PCDATA)>

]>
```

2.4. 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:string"/>
<xs:element name="translator" type="xs:string"/>
<xs:element name="publisher" type="xs:string"/>
</xs:sequence>
</xs:complexType>
</xs:element>
```


</xs: schema>