

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

Sprawozdanie 3:
Resource Description Framework (RDF) in use

Krzysztof Kutt
Michał Nowak

Podstawy Inżynierii Wiedzy
Środa 14:00

Kraków 2012

1 Semantic vocabularies: Dublin Core

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE rdf:RDF PUBLIC "-//DUBLIN CORE//DCMES DTD 2002/07/31//EN" "http://
/dublincore.org/documents/2002/07/31/dcmes-xml/dcmes-xml-dtd.dtd">
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns="http://student.agh.edu.pl/~kkutt/piw2012/mylibrary#">
<rdf:Description rdf:about="http://student.agh.edu.pl/~kkutt/piw2012/
mylibrary#book-czystykod">
  <author rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Robert C.
Martin</author>
  <dc:title rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Czysty
kod. Pordręcznik dobrego programisty</dc:title>
<dc:publisher rdf:datatype="http://www.w3.org/2001/
/XMLSchema#string">Helion</dc:publisher>
  <dc:language>pl</dc:language>
</rdf:Description>
<rdf:Description rdf:about="http://student.agh.edu.pl/~kkutt/piw2012/
mylibrary#book-wprowadzeniedoalg">
  <authors rdf:parseType="Collection">
    <rdf:Description>
      <rdf:first rdf:datatype="http://www.w3.org/2001/
/XMLSchema#string">Thomas H. Cormen</rdf:first>
      <rdf:rest> <rdf:Description>
        <rdf:first rdf:datatype="http://www.w3.org/2001/
/XMLSchema#string">Charles E. Leiserson</rdf:first>
        <rdf:rest> <rdf:Description>
          <rdf:first rdf:datatype="http://www.w3.org/2001/
/XMLSchema#string">Ronald L. Rivest</rdf:first>
          <rdf:rest> <rdf:Description>
            <rdf:first rdf:datatype="http://www.w3.org/2001/
/XMLSchema#string">Clifford Stein</rdf:first>
            <rdf:rest rdf:resource="http://www.w3.org/1999/02/22-rdf-
syntax-ns#nil" />
          </rdf:Description> </rdf:rest>
        </rdf:Description> </rdf:rest>
      </rdf:Description> </rdf:rest>
    </authors>
    <dc:title rdf:datatype="http://www.w3.org/2001/
/XMLSchema#string">Wprowadzenie do algorytmów</dc:title>
    <dc:publisher rdf:datatype="http://www.w3.org/2001/
/XMLSchema#string">Wydawnictwa Naukowo-Techniczne</dc:publisher>
    <dc:language>pl</dc:language>
  </rdf:Description>
<rdf:Description rdf:about="http://student.agh.edu.pl/~kkutt/piw2012/
mylibrary#book-sztukapodstepu">
  <author rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Kevin
Mitnick</author>
  <author rdf:datatype="http://www.w3.org/2001/XMLSchema#string">William
```

```

Simon</author>
  <dc:title rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Sztuka
  podstepu</dc:title>
  <dc:publisher rdf:datatype="http://www.w3.org/2001/
  XMLSchema#string">Helion</dc:publisher>
  <dc:language>pl</dc:language>
</rdf:Description>
<rdf:Description rdf:about="http://student.agh.edu.pl/~kkutt/piw2012/
mylibrary#book-wprowadzeniedoteoriigrafow">
  <author rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Robin J.
  Wilson</author>
  <dc:title rdf:datatype="http://www.w3.org/2001/
  XMLSchema#string">Wprowadzenie do teorii grafów</dc:title>
  <dc:publisher rdf:datatype="http://www.w3.org/2001/
  XMLSchema#string">Wydawnictwo Naukowe PWN</dc:publisher>
  <dc:language>pl</dc:language>
</rdf:Description>
<rdf:Description rdf:about="http://student.agh.edu.pl/~kkutt/piw2012/
mylibrary#book-jakpisacpraczezpsycho">
  <author rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Edward
  Nęcka</author>
  <author rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Ryszard
  Stocki</author>
  <dc:title rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Jaki
  pisać prace z psychologii</dc:title>
  <dc:publisher rdf:datatype="http://www.w3.org/2001/
  XMLSchema#string">Universitas</dc:publisher>
  <dc:language>pl</dc:language>
</rdf:Description>
<rdf:Description rdf:about="http://student.agh.edu.pl/~kkutt/piw2012/
mylibrary#cd-iananderson_playsorchestraljethrotull">
  <artist rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Ian
  Anderson</artist>
  <dc:title rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Plays
  Orchestral Jethro Tull</dc:title>
  <dc:language>en</dc:language>
</rdf:Description>
<rdf:Description rdf:about="http://student.agh.edu.pl/~kkutt/piw2012/
mylibrary#cd-rhcp_stadiumarcadium">
  <artist rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Red Hot
  Chilli Peppers</artist>
  <dc:title rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Stadium
  Arcadium</dc:title>
  <dc:language>en</dc:language>
</rdf:Description>

<rdf:Description rdf:about="http://student.agh.edu.pl/~kkutt/piw2012/
mylibrary">
  <MyFavouriteBooks>
    <rdf:Bag>
      <rdf:li rdf:resource="http://student.agh.edu.pl/~kkutt/piw2012/
mylibrary#book-wprowadzeniedoalg"/>
      <rdf:li rdf:resource="http://student.agh.edu.pl/~kkutt/piw2012/

```

```

mylibrary#book-wprowadzeniedoteoriigrafow"/>
  </rdf:Bag>
</MyFavouriteBooks>
</rdf:Description>
</rdf:RDF>

```

2 RDFSchema

```

<?xml version="1.0" encoding="UTF-8"?>
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xml:base="http://student.agh.edu.pl/~kkutt/piw2012/mylibrary#">

<rdfs:Class rdf:ID="MultimediaItem" />
<rdfs:Class rdf:ID="MusicCD">
  <rdfs:subClassOf rdf:resource="#MultimediaItem"/>
  <rdfs:label>Klasa muzycznych CD</rdfs:label>
</rdfs:Class>
<rdfs:Class rdf:ID="Book">
  <rdfs:subClassOf rdf:resource="#MultimediaItem"/>
  <rdfs:label>Klasa książek</rdfs:label>
  <rdfs:comment>Klasa wszystkich książek z biblioteczki</rdfs:comment>
</rdfs:Class>

<rdf:Description rdf:about="http://student.agh.edu.pl/~kkutt/piw2012/
mylibrary#book-czystykod">
  <author rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Robert C.
Martin</author>
  <title rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Czysty kod.
Pordęcznik dobrego programisty</title>
<publisher rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Helion</
publisher>
  <language>pl</language>
  <rdf:type rdf:resource="http://student.agh.edu.pl/~kkutt/piw2012/
mylibrary#Book"/>
</rdf:Description>

<rdf:Description rdf:about="http://student.agh.edu.pl/~kkutt/piw2012/
mylibrary#book-wprowadzeniedoalg">
  <authors rdf:parseType="Collection">
  <rdf:Description>
    <rdf:first rdf:datatype="http://www.w3.org/2001/
XMLSchema#string">Thomas H. Cormen</rdf:first>
    <rdf:rest> <rdf:Description>
      <rdf:first rdf:datatype="http://www.w3.org/2001/
XMLSchema#string">Charles E. Leiserson</rdf:first>
      <rdf:rest> <rdf:Description>
        <rdf:first rdf:datatype="http://www.w3.org/2001/
XMLSchema#string">Ronald L. Rivest</rdf:first>
        <rdf:rest> <rdf:Description>
          <rdf:first rdf:datatype="http://www.w3.org/2001/

```

```
XMLSchema#string">Clifford Stein</rdf:first>
    <rdf:rest rdf:resource="http://www.w3.org/1999/02/22-rdf-
syntax-ns#nil" />
    </rdf:Description> </rdf:rest>
    </rdf:Description> </rdf:rest>
    </rdf:Description> </rdf:rest>
</rdf:Description>
</authors>
<title rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Wprowadzenie
do algorytmów</title>
<publisher rdf:datatype="http://www.w3.org/2001/
XMLSchema#string">Wydawnictwa Naukowo-Techniczne</publisher>
    <language>pl</language>
    <rdf:type rdf:resource="http://student.agh.edu.pl/~kkutt/piw2012/
mylibrary#Book"/>
</rdf:Description>
<rdf:Description rdf:about="http://student.agh.edu.pl/~kkutt/piw2012/
mylibrary#book-sztukapodstepu">
    <author rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Kevin
Mitnick</author>
    <author rdf:datatype="http://www.w3.org/2001/XMLSchema#string">William
Simon</author>
    <title rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Sztuka
podstepu</title>
<publisher rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Helion</
publisher>
    <language>pl</language>
    <rdf:type rdf:resource="http://student.agh.edu.pl/~kkutt/piw2012/
mylibrary#Book"/>
</rdf:Description>
<rdf:Description rdf:about="http://student.agh.edu.pl/~kkutt/piw2012/
mylibrary#book-wprowadzeniedoteoriigrafow">
    <author rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Robin J.
Wilson</author>
<title rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Wprowadzenie
do teorii grafów</title>
<publisher rdf:datatype="http://www.w3.org/2001/
XMLSchema#string">Wydawnictwo Naukowe PWN</publisher>
    <language>pl</language>
    <rdf:type rdf:resource="http://student.agh.edu.pl/~kkutt/piw2012/
mylibrary#Book"/>
</rdf:Description>
<rdf:Description rdf:about="http://student.agh.edu.pl/~kkutt/piw2012/
mylibrary#book-jakpisacpracezpsycho">
    <author rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Edward
Nęcka</author>
    <author rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Ryszard
Stocki</author>
    <title rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Jaki pisać
prace z psychologii</title>
<publisher rdf:datatype="http://www.w3.org/2001/
XMLSchema#string">Universitas</publisher>
    <language>pl</language>
```

```

    <rdf:type rdf:resource="http://student.agh.edu.pl/~kkutt/piw2012/
mylibrary#Book"/>
</rdf:Description>
<rdf:Description rdf:about="http://student.agh.edu.pl/~kkutt/piw2012/
mylibrary#cd-iananderson_playsorchestraljethrotull">
    <artist rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Ian
Anderson</artist>
    <title rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Plays
Orchestral Jethro Tull</title>
    <language>en</language>
    <rdf:type rdf:resource="http://student.agh.edu.pl/~kkutt/piw2012/
mylibrary#MusicCD"/>
</rdf:Description>
<rdf:Description rdf:about="http://student.agh.edu.pl/~kkutt/piw2012/
mylibrary#cd-rhcp_stadiumarcadium">
    <artist rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Red Hot
Chilli Peppers</artist>
    <title rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Stadium
Arcadium</title>
    <rdf:type rdf:resource="http://student.agh.edu.pl/~kkutt/piw2012/
mylibrary#MusicCD"/>
    <language>en</language>
</rdf:Description>

<rdf:Description rdf:about="http://student.agh.edu.pl/~kkutt/piw2012/
mylibrary">
    <MyFavouriteBooks>
        <rdf:Bag>
            <rdf:li rdf:resource="http://student.agh.edu.pl/~kkutt/piw2012/
mylibrary#book-wprowadzeniedoalg"/>
            <rdf:li rdf:resource="http://student.agh.edu.pl/~kkutt/piw2012/
mylibrary#book-wprowadzeniedoteoriigrafow"/>
        </rdf:Bag>
    </MyFavouriteBooks>
</rdf:Description>
</rdf:RDF>

```

3 SPARQL - demo

Uruchomiliśmy wszystkie przykładowe zapytania. Jednak wg nas występują jakieś problemy z repozytoriami danych, ponieważ zwracają one w wynikach puste tabele. Tylko pierwsze (domyślne) zapytanie zwróciło wartości: dla jednego repozytorium 1 wynik, dla drugiego 2 wyniki.

- **What semantic vocabularies are used in the queries? What are they for?**

W zapytaniach wykorzystywane są dwa typy słów kluczowych:

- nazwy atrybutów z odpowiednich przestrzeni nazw, np `name`, `nick` z *FOAF*, czy `title`, `creator` z *Dublin Core*
- słowa kluczowe języka *SPARQL*: `SELECT`, `DISTINCT`, `ORDER BY`, `WHERE`,

LIMIT (wszystkie znane z języka *SQL*), PREFIX (definiujący przestrzenie nazw, analogicznie jak w *RDF*)

- **What do SELECT queries do?**
Pobierają odpowiednie informacje i zwracają je w formie tabelarycznej
- **What do CONSTRUCT queries do?**
Pobierają odpowiednie informacje i zwracają je w formie poprawnego pliku RDF

4 SPARQL queries - basics

- **friends who have name and e-mail defined**

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
```

```
SELECT DISTINCT ?name ?mbox
WHERE {
    ?x rdf:type foaf:Person .
    ?x foaf:name ?name.
    ?x foaf:mbox ?mbox
}
```

- **friends who have name and e-mail defined and optional homepage**

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
```

```
SELECT DISTINCT ?name ?mbox ?homepage
WHERE {
    ?x rdf:type foaf:Person .
    ?x foaf:name ?name.
    ?x foaf:mbox ?mbox.
    OPTIONAL{?x foaf:homepage ?homepage}
}
```

- **friends who have name and e-mail defined and optional homepage, sorted by name descending**

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
```

```
SELECT DISTINCT ?name ?mbox ?homepage
WHERE {
    ?x rdf:type foaf:Person .
    ?x foaf:name ?name.
    ?x foaf:mbox ?mbox.
    OPTIONAL{?x foaf:homepage ?homepage}
}
ORDER BY DESC(?name)
```

5 SPARQL queries - options

- **people whose name starts with 'K'**

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
```

```
SELECT DISTINCT ?name
WHERE {
    ?x rdf:type foaf:Person .
    ?x foaf:name ?name.
    FILTER regex(?name, "^K")
}
```

- **people who are older than 18 years old**

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
```

```
SELECT DISTINCT ?name ?age
WHERE {
    ?x rdf:type foaf:Person .
    ?x foaf:name ?name.
    ?x foaf:age ?age.
    FILTER (?age > '18')
}
```

- **people whose name starts with 'K' or are older than 18 years old, make search caseinsensitive**

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
```

```
SELECT DISTINCT ?name ?age
WHERE {
    ?x rdf:type foaf:Person .
    {
        ?x foaf:name ?name.
        FILTER regex(?name, "^K","i")
    } UNION {
        ?x foaf:name ?name.
        ?x foaf:age ?age.
        FILTER (?age > '18')
    }
}
```

- **people having e-mails on student.agh.edu.pl server**

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
```

```
SELECT DISTINCT ?name ?mbox
WHERE {
    ?x rdf:type foaf:Person .
    ?x foaf:name ?name.
```



```

    ?x foaf:mbox ?mbox.
  FILTER regex(?mbox, "@student.agh.edu.pl$")
}

```

- **name of people, who have homepage or e-mail on student.agh.edu.pl server**

```

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>

```

```

SELECT DISTINCT ?name ?homepage ?mbox
WHERE {
  ?x rdf:type foaf:Person .
  {
    ?x foaf:name ?name.
    ?x foaf:homepage ?homepage
    FILTER regex(?homepage, "^http://student.agh.edu.pl")
  } UNION {
    ?x foaf:name ?name.
    ?x foaf:mbox ?mbox.
    FILTER regex(?mbox, "@student.agh.edu.pl$")
  }
}

```

6 Open Data Sets

Głównym ograniczeniem w korzystaniu z tych zbiorów danych są długie czasy wykonywania zapytań. Wynika to z ogromnej ilości danych, na których operuje wyszukiwanie. Gdy zapytanie jest skomplikowane, czas działania jest jeszcze dłuższy.