

Język RDF

Mikołaj Morzy
Agnieszka Ławrynowicz

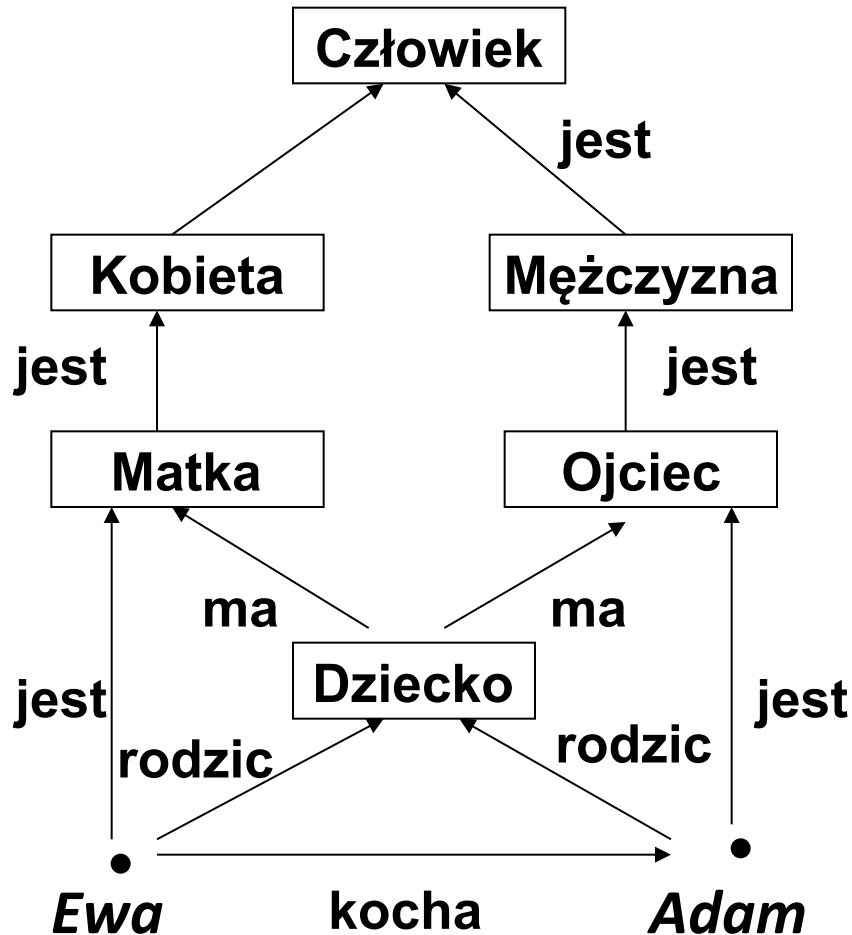
Instytut Informatyki
Poznań, rok akademicki 2013/2014

Sieci semantyczne

sieci semantyczne = graficzna notacja do reprezentacji **wiedzy** reprezentowanej jako zbiór **węzłów (pojęć)** połączonych etykietowanymi **łukami**, które reprezentują **relacje** między węzłami

- w ramach sztucznej inteligencji - początki w latach 60tych XX wieku
- dużo wcześniej wykorzystywane w filozofii, psychologii, lingwistyce

Sieci semantyczne - przykład



Ewa **jest** **matką**

Ewa **kocha** *Adama*

Każde **dziecko** **ma** **ojca**

Każda **matka** **jest** **kobietą**

Każda **kobieta** **jest** **człowiekiem**

...

„Trójkowy” model danych

RDF = Resource Description Framework

<http://www.w3.org/RDF/>

- **Zdania w postaci trójek (ang. *triples*):**
<subject, predicate, object>
 - **Podmiot (ang. Subject):** zasób lub węzeł anonimowy (ang. *blank node*)
 - **Predykat (ang. Predicate):** własność
 - **Obiekt (ang. Object):** zasób, literał lub węzeł anonimowy
- **Przykład:**
<ex:Ewa,ex:kocha,ex:Adam>

Podstawowe elementy

- Zasoby (**ang. *resources***)
 - *identyfikowane* za pomocą URI, ale ...URI niekoniecznie *wskazuje* zasób
 - odpowiadają węzłom w grafie
 - np. <http://przyklad.org/#Anna>
- Własności (**ang. *properties***)
 - *identyfikowane* za pomocą URI
 - odpowiadają łukom w grafie
 - relacja binarna między dwoma zasobami
 - np. <http://przyklad.org/#jestWlascicielem>
- Literały (**ang. *literals***)
 - konkretne wartości danych np. „VW Passat”, "2009-09-29"

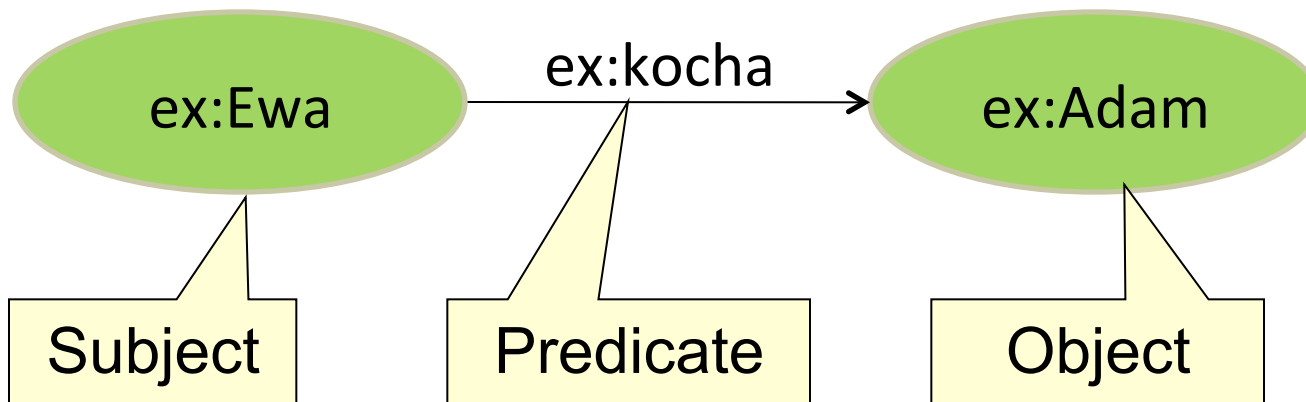
URI = Universal Resource Identifier

“An URI is a compact string of characters for identifying an abstract or physical resource” (RFC 2396)

- URI mogą ale nie muszą odnosić się do fizycznych zasobów:
 - **URL (Uniform Resource Locator):** specyficzny typ URI, zasoby dostępne w sieci WWW
<http://www.cs.put.poznan.pl>
 - **URN (Uniform Resource Name):** specyfikują nazwę zasobu, niezależnie od jego lokalizacji
<urn:issn:1560-1560>
- W przypadku RDF, URI zazwyczaj wygląda jak ‘normalny’ URL, często z **identyfikatorem fragmentu** (po znaku #)
<http://dublincore.org/usage/documents/principles/#element>

Trójka jako graf

- **Etykietowane grafy skierowane**
 - Węzły: zasoby, literały
 - Etykiety: własności
 - Łuki: zdania



Zasoby

- **Zasobem może być:**
 - strona WWW (np. <http://www.w3.org>)
 - osoba (np. <http://www.cs.put.poznan.pl/alawrynowicz>)
 - książka (np. isbn:0130273635)
 - cokolwiek wskazanego przez URI
- **URI jest *identifikatorem* zasobu, a niekoniecznie wskazuje jego lokalizację w sieci WWW**
- **Zdania na temat zasobów:**

<http://www.w3.org> **ma format** text/html

Literały

- **Proste**
 - np. `"napis"`
- **Typowane**
 - np. `"czesc"^^xsd:string, "100"^^xsd:integer`
 - rekomendowane typy danych: typy danych XML Schema
 - możliwość rozszerzenia typów
- **Tylko jako *object* trójki.:**
 - `<http://przyklad.org/#Anna>`,
 - `<http://przyklad.org/#nazywaSie>`,
 - `"Anna Nowak"^^xsd:string`

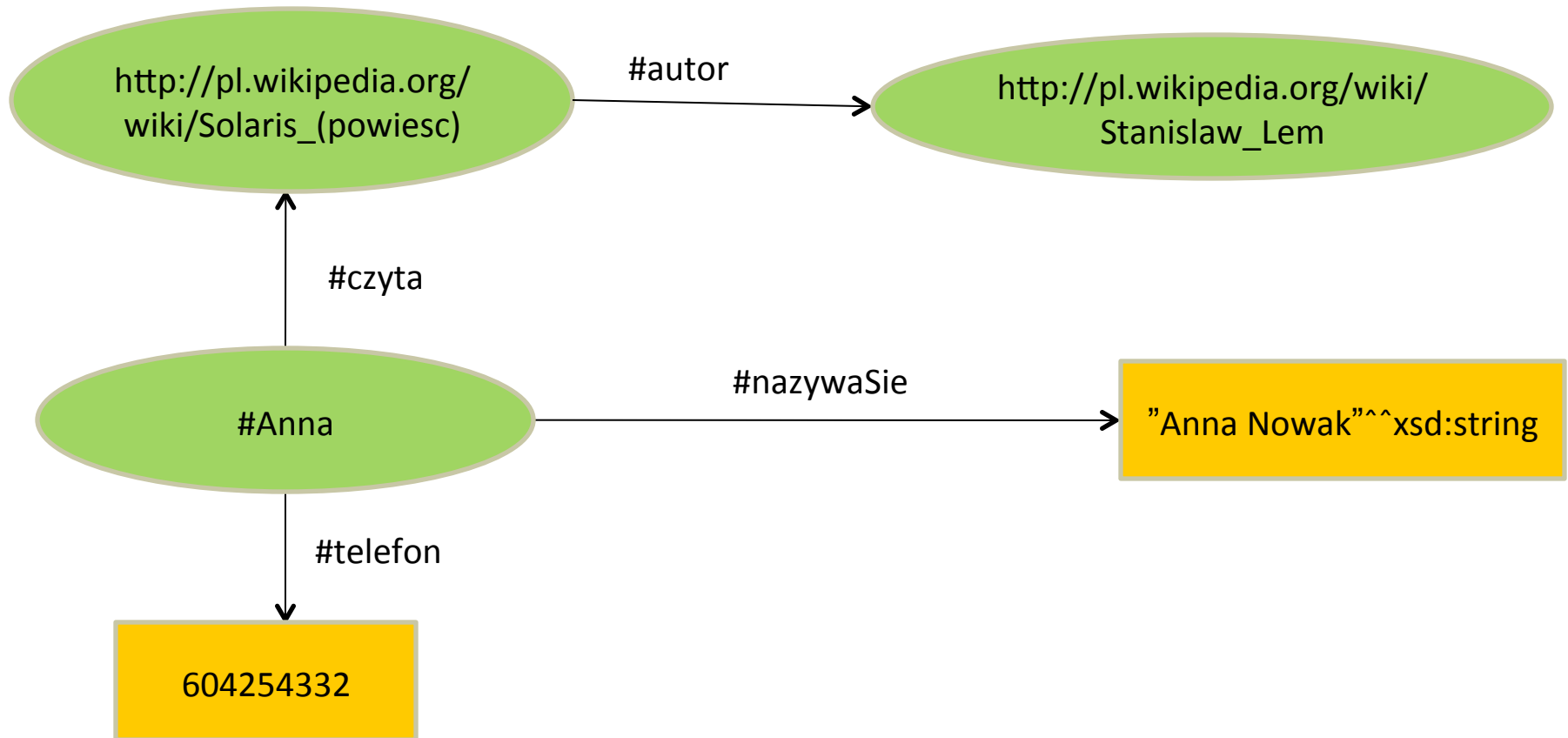
`http://przyklad.org/#Anna`

`http://przyklad.org/#nazywaSie`

`"Anna Nowak"^^xsd:string`

Łączenie trójek

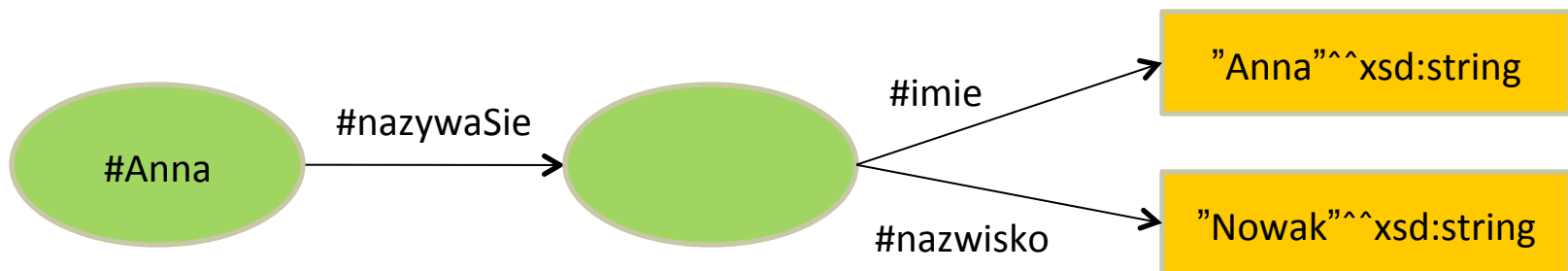
Sieć semantyczna



Węzły anonimowe

- Węzły anonimowe (ang. *blank nodes*) nie posiadają URI
 - nienazwane zasoby
 - bardziej złożone konstrukcje
- Reprezentacja węzłów anonimowych jest zależna od *składni*
- Przykład:

```
<#Anna>,<#nazywaSie>, _:imieNazwiskoAnny>  
<_:imieNazwiskoAnny,<#imie>,"Anna"^^xsd:string>  
<_:imieNazwiskoAnny,<#nazwisko>,"Nowak"^^xsd:string>
```



Reifikacja

- zdania o zdaniach

Marek twierdzi, że Anna nazywa się "Anna Nowak"

```
<#mojeZdanie>,rdf:type,rdf:Statement>
```

```
<#mojeZdanie>,rdf:subject,<#Anna>
```

```
<#mojeZdanie>,rdf:predicate,<#nazywaSie>
```

```
<#mojeZdanie>,rdf:object,"Anna Nowak">
```

```
<#Marek>,<#twierdzi>,<#mojeZdanie>
```

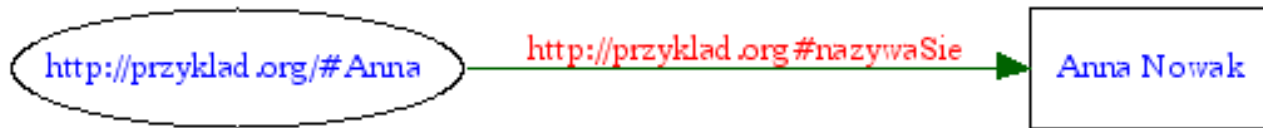
Słownictwo

- **Słownictwo RDF jest zdefiniowane w przestrzeni nazw:**
<http://www.w3.org/1999/02/22-rdf-syntax-ns#>
- **Klasy:**
rdf:Property rdf:Statement rdf:XMLLiteral rdf:Seq rdf:Bag
rdf:Alt rdf:List
- **Własności:**
rdf:type rdf:subject rdf:predicate rdf:object rdf:first rdf:rest
rdf:n rdf:value
- **Zasoby:**
rdf:nil

Serializacja RDF do XML

- Dokument RDF może być reprezentowany w XML za pomocą znacznika `rdf:RDF`
- Dokument XML w tym przypadku zawiera elementy opatrzone znacznikiem `rdf:Description`
- Każdy taki znacznik zawiera zdanie na temat zasobu identyfikowane za pomocą trzech sposobów:
 - atrybutu `about` – odnoszącego się do istniejącego zasobu
 - atrybutu typu `ID` – tworzącego nowy zasób
 - bez nazwy – tworzącego zasób anonimowy

Serializacja do XML – przykład 1

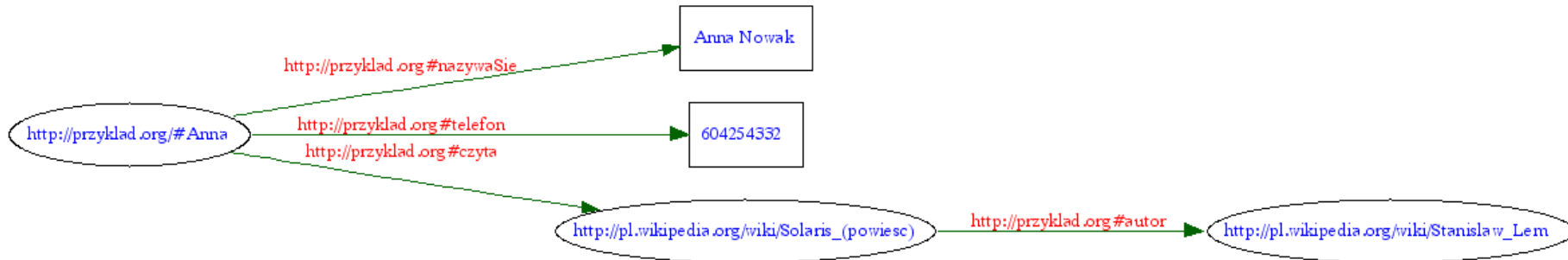


```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:ex="http://przyklad.org/#">

  <rdf:Description rdf:about="http://przyklad.org/#Anna">
    <ex:nazywaSie>Anna Nowak</ex:nazywaSie>
  </rdf:Description>

</rdf:RDF>
```

Serializacja do XML – przykład 2

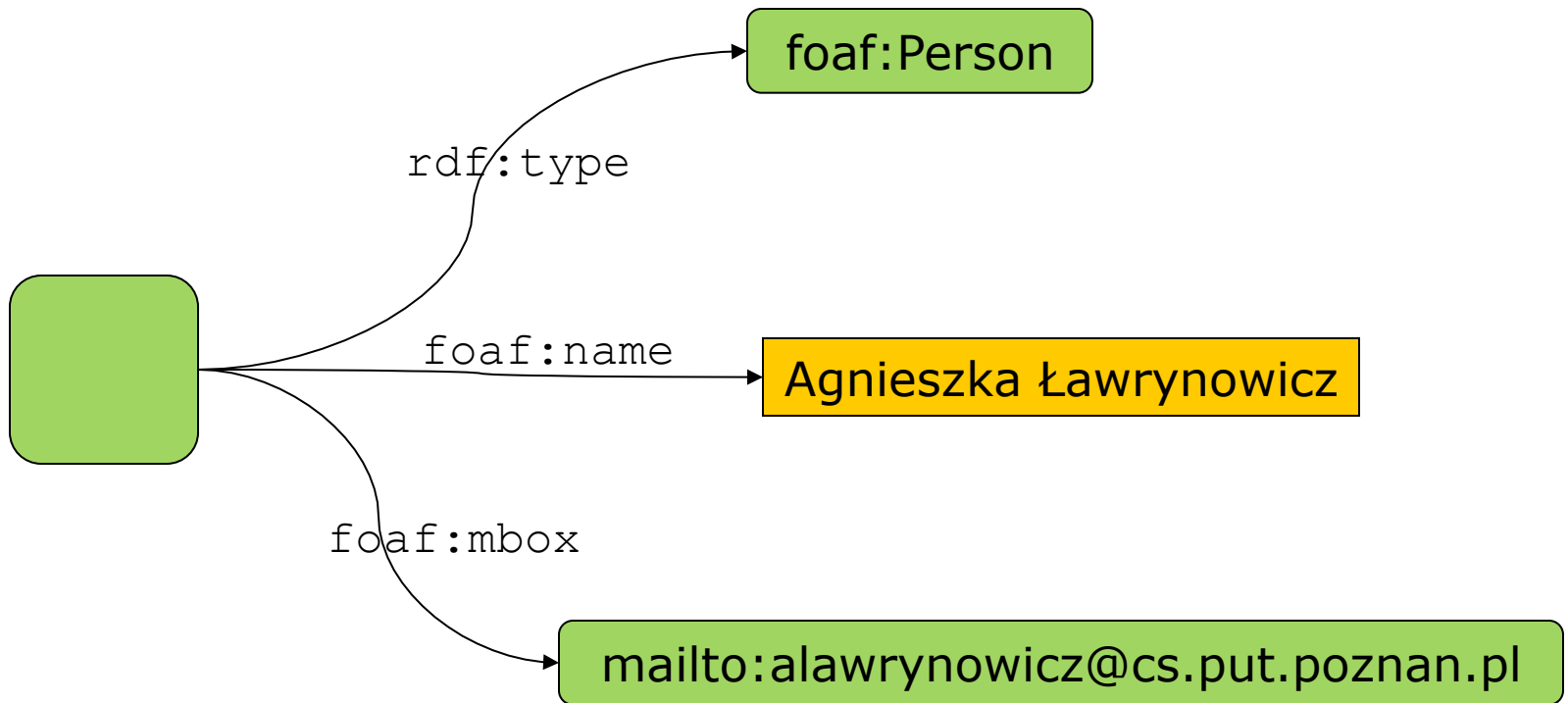


```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
  xmlns:ex="http://przyklad.org#">

  <rdf:Description rdf:about="http://przyklad.org/#Anna">
    <ex:nazywaSie rdf:datatype="xsd:string">Anna Nowak</ex:nazywaSie>
    <ex:telefon rdf:datatype="xsd:integer">604254332</ex:telefon>
    <ex:czyta rdf:resource="http://pl.wikipedia.org/wiki/Solaris_(powiesc)"/>
  </rdf:Description>

  <rdf:Description rdf:about="http://pl.wikipedia.org/wiki/Solaris_(powiesc)">
    <ex:autor rdf:resource="http://pl.wikipedia.org/wiki/Stanislaw_Lem"/>
  </rdf:Description>
</rdf:RDF>
```


Węzeł anonimowy – przykład



Węzeł anonimowy – serializacja do XML

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

  <foaf:Person>
    <foaf:name>Agnieszka Ławrynowicz</foaf:name>
    <foaf:mbox rdf:resource="mailto:alawrynowicz@cs.put.poznan.pl" />
  </foaf:Person>

</rdf:RDF>
```