

Szemantikus technológiák alapjai

(RDF, Ontológiák, SPARQL)

Nemzeti Fejlesztési Ügynökség
www.ujszechenyiterv.gov.hu
06 40 638 638



The Project is supported by the European Union and co-financed by the European Social Fund (grant agreement no. TAMOP 4.2.1./B-09/1/KMR-2010-0003).

Pinczel Balázs, 2011. 10. 14.

Szemantikus web

- rengeteg publikus adat a weben :))
- forma: általában szöveges :((
 - emberek által jól érthető, de az **automatikus feldolgozás nehéz**
- szemantikus web célja: ez a tudás gépileg feldolgozható legyen

Motiváció

- példa: keresés "hotels near Hyde Park"
- régebben ez szöveges keresés volt
- szemantikát is figyelembe véve:
 - keresésben "Hyde Park"
→ koordináták
 - talált oldalakon cím
→ koordináták
 - koordináták összehasonlítása



RDF állítások

- RDF: Resource Description Framework
- állítások **alany-állítmány-tárgy** alakban
- mindháromhoz egy-egy IRI-t rendelünk
- példa: „természetes nyelvi” szöveg

```
http://www.example.org/index.html has a creator  
whose value is John Smith
```

- RDF hármás (triple):

```
<http://www.example.org/index.html>  
<http://purl.org/dc/elements/1.1/creator>  
<http://www.example.org/staffid/85740> .
```

Rövidítés, prefix

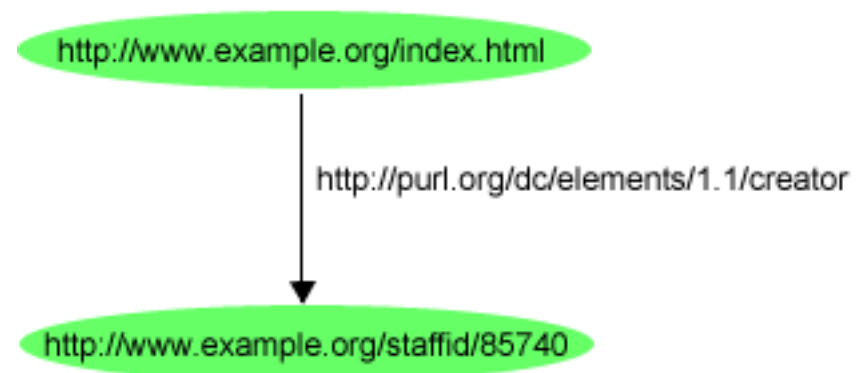
- az IRI-k gyakran ugyanúgy kezdődnek
- az ismétlődő prefixekre bevezethetünk rövidítést *(csak egyszer kell)*
- példa: az előző hármassal ekvivalens

```
@prefix ex:      <http://www.example.org/> .  
@prefix dc:      <http://purl.org/dc/elements/1.1/> .  
@prefix exstaff: <http://www.example.org/staffid/> .  
  
ex:index.html dc:creator exstaff:85740 .
```

Gráf modell

- a hármások felfoghatók irányított élekként:
 - alany (subject): az él kezdőpontjának címkéje
 - állítmány (predicate): az él címkéje
 - tárgy (object): az él végpontjának címkéje
- példa:

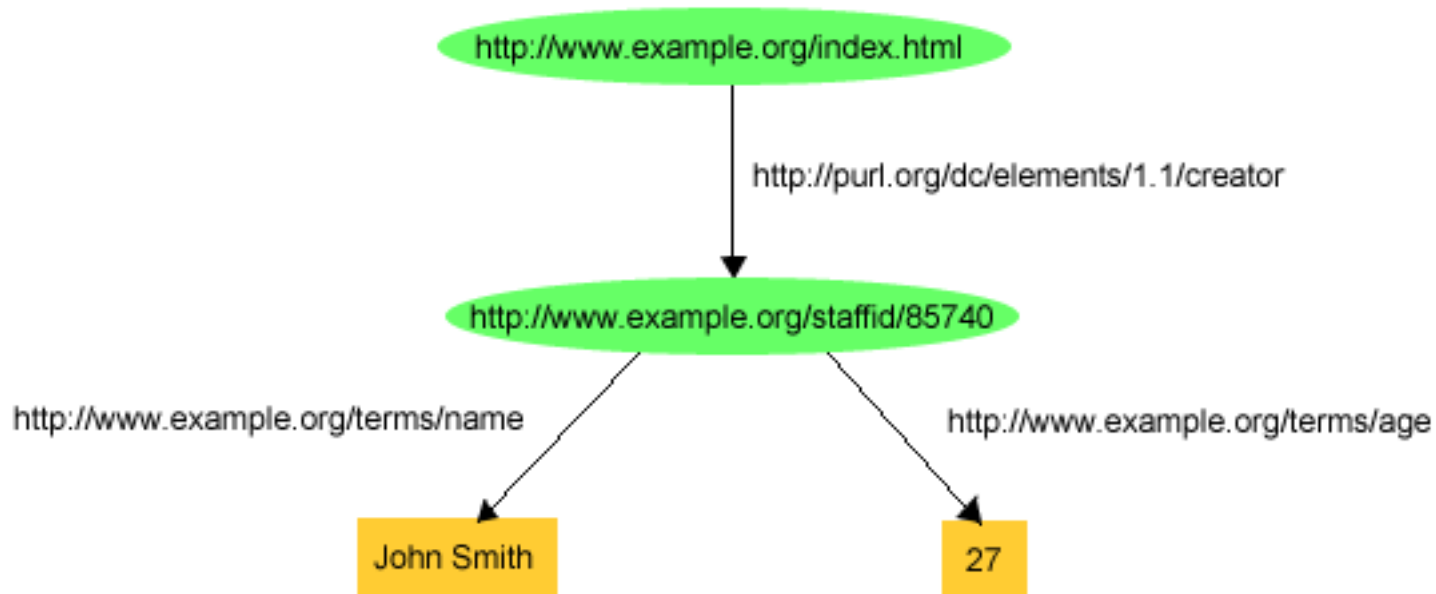
```
ex:index.html dc:creator exstaff:85740 .
```



Literálok

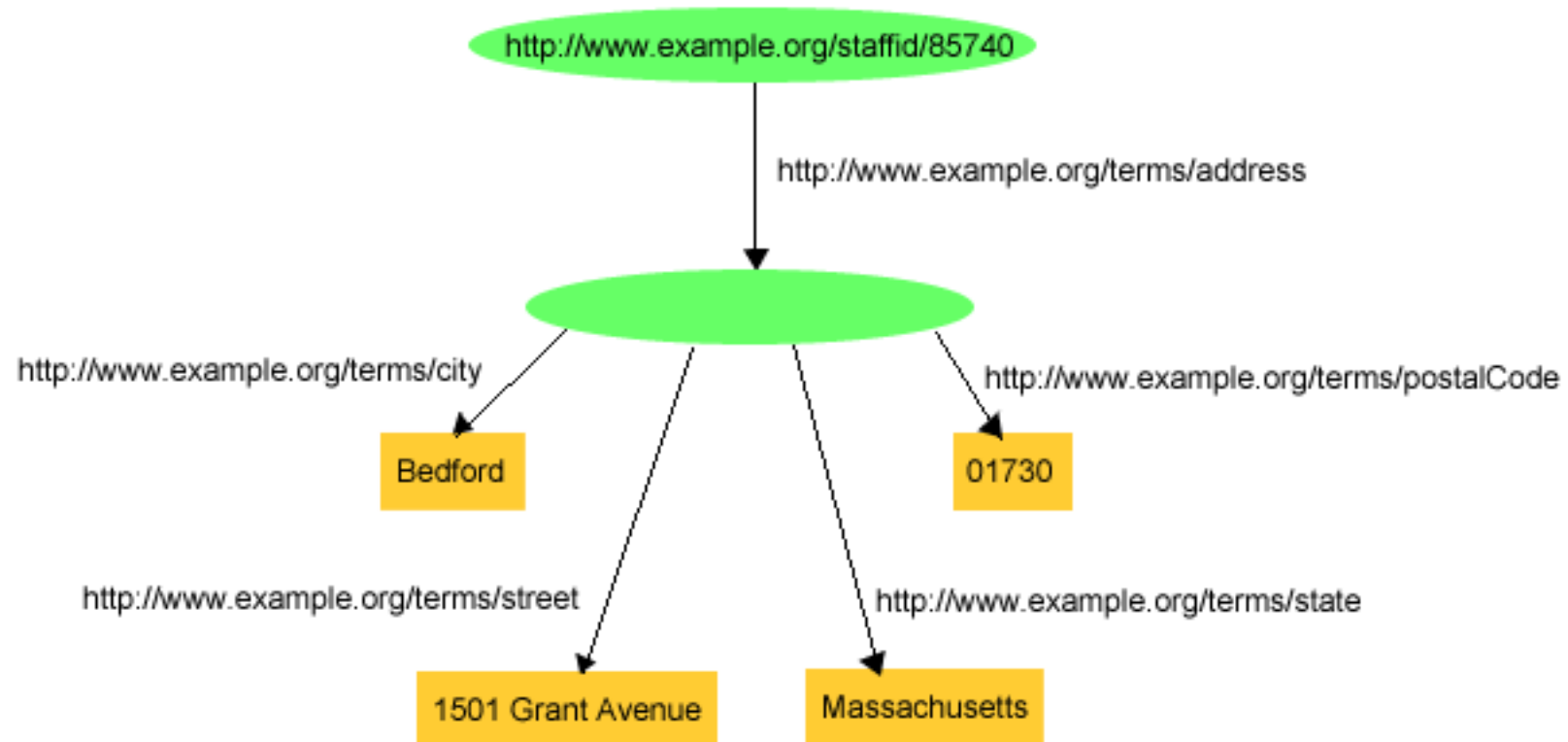
- a hármásokban a tárgy helyén lehetnek:

```
exstaff:85740 exterms:name "John Smith" .  
exstaff:85740 exterms:age "27" .
```



Üres csúcsok

- üres csúcs (blank node) az alany, vagy a tárgy helyén lehet
- összefog több csúcsot, névtelenül



Üres csúcs leírása

- a "_" karakterrel, mint prefix-szel, utána azonosító
- példa: az előző gráf szövegesen

```
exstaff:85740    exterms:address    _:johnaddress .  
_:johnaddress  exterms:street     "1501 Grant Avenue" .  
_:johnaddress  exterms:city       "Bedford" .  
_:johnaddress  exterms:state      "Massachusetts" .  
_:johnaddress  exterms:postalCode "01730" .
```

RDF formátumok

- **Notation3 \supseteq Turtle \supseteq N-Triples**
 - szabad szemmel is átlátható
 - a bővebbekben szintaktikus egyszerűsítések
 - (az itteni példák Turtle-ben vannak)
- **RDF/XML, példa:**

```
<?xml version="1.0"?>
<rdf:RDF xmlns:ex="..." ... >
  <rdf:Description rdf:about="http://www.example.org/index.html">
    <dc:creator rdf:resource="http://www.example.org/staffid/85740"
      exterms:name="John Smith"
      exterms:age="27" />
  </rdf:Description>
</rdf:RDF>
```

Szókészlet, ontológia

- gépi feldolgozhatósághoz kéne: ugyanazt a dolgot ugyanúgy hívjuk
- megoldás: szókészletek, ontológiák
- szókészlet (vocabulary): megfeleltetés a valós világ fogalmai és IRI-k között
- **ontológia**: szókészlet + a definiált fogalmak közötti kapcsolatok, megszorítások, és szabályok

Szókészlet példa

- **Dublin Core** (dc): erőforrások metaadatai
- **Friend of a Friend** (foaf): közösségi hálók
- **RDF**: néhány „beépített” fogalom

```
ex:index.html dc:creator exstaff:85740 .  
exstaff:85740 rdf:type foaf:Person .  
exstaff:85740 foaf:name "John Smith" .
```

- szókészlet, vagy ontológia definiálása:
szintén RDF-ben!

RDF Séma

- szókészlet saját szókészletek leírására

```
ex:Person rdf:type          rdfs:Class .
ex:Female rdfs:subClassOf  ex:Person .
ex:Male   rdfs:subClassOf  ex:Person .
ex:name   rdf:type          rdf:Property .
ex:name   rdfs:domain       ex:Person .
```

```
exstaff:85740 ex:name "John Smith" .
```

- kikövetkeztethető (többek között):

```
exstaff:85740 rdf:type ex:Person .
```

OWL

- **Web Ontology Language**, ontológiák leírására való **RDF alapú** nyelv
- három (egyre bővülő) résznyelv, egyre nagyobb kifejezőerővel
- OWL Lite: hierarchiák, megszorítások
- OWL DL: maximális kifejezőerő az eldönthetőségén belül
- OWL Full: maximális kifejezőerő

OWL példák

- ha OWL-al leírjuk, hogy az `ex:superiorOf` tulajdonság tranzitív, akkor felhasználható következtetésben:

```
exstaff:john ex:superiorOf exstaff:peter .  
exstaff:peter ex:superiorOf exstaff:tom .
```



```
exstaff:john ex:superiorOf exstaff:tom .
```

- ha OWL-al leírjuk, hogy az `ex:parentOf` tulajdonság inverze az `ex:childOf`:

```
ex:mary ex:parentOf ex:kate .
```



```
ex:kate ex:childOf ex:mary .
```

SPARQL

- az RDF lekérdezőnyelve
- SQL-hez hasonló záradékok
- egy gráfmintára illeszkedő részeket keres a gráfban

SPARQL példa

- **adathalmaz:**

```
_:a foaf:name "Johnny Lee Outlaw" .
_:a foaf:mbox <mailto:jlow@example.com> .
_:b foaf:name "Peter Goodguy" .
_:b foaf:mbox <mailto:peter@example.org> .
_:c foaf:name "John Smith" .
```

- **lekérdezés:**

```
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
SELECT ?name ?mbox
WHERE
  { ?x foaf:name ?name .
    ?x foaf:mbox ?mbox }
```

- **eredmény:**

name	mbox
"Johnny Lee Outlaw"	<mailto:jlow@example.com>
"Peter Goodguy"	<mailto:peter@example.org>

Opcionális illesztés

- lekérdezés:

```
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
SELECT ?name ?mbox
WHERE
{
  ?x foaf:name ?name .
  OPTIONAL { ?x foaf:mbox ?mbox }
}
```

- eredmény:

name	mbox
"Johnny Lee Outlaw"	<mailto:jlow@example.com>
"Peter Goodguy"	<mailto:peter@example.org>
"John Smith"	

Távoli adathalmaz lekérdezése

- lokális és távoli adatok elérése egy lekérdezésen belül

```
PREFIX exterm: <http://www.example.org/terms/>
PREFIX foaf:   <http://xmlns.com/foaf/0.1/>
PREFIX dbpedia: <http://dbpedia.org/ontology/>

SELECT ?name ?country ?capital
WHERE
{
  ?x foaf:name ?name .
  ?x exterm:livesIn ?country .
  SERVICE <http://dbpedia.org/sparql>
    { ?country dbpedia:capital ?capital }
}
```

Szűrés

- példa:

```
PREFIX exterm: <http://www.example.org/terms/>
PREFIX foaf:   <http://xmlns.com/foaf/0.1/>
SELECT ?name ?age
WHERE
{
  ?x foaf:name ?name .
  ?x exterm:age ?age .
  FILTER (?age < 30)
}
```

Egyéb lehetőségek

- mint az SQL-ben:
 - **ORDER BY**
 - **LIMIT**
 - **OFFSET**
 - **DISTINCT**

Más lekérdezéstípusok

- **ASK:** van-e a mintának megfelelő részgráf

```
PREFIX foaf: <http://xmlns.com/foaf/0.1/>  
ASK { ?x foaf:name "Alice" }
```

(eredménye „yes” vagy „no”)

- **CONSTRUCT:** gráfépítés sablon alapján

```
PREFIX ex: <http://www.example.org/>  
PREFIX foaf: <http://xmlns.com/foaf/0.1/>  
CONSTRUCT { ex:John foaf:knows ?x }  
WHERE { ?x foaf:name "Alice" }
```

- van még: DESCRIBE
- lesz: INSERT, DELETE, CREATE, DROP, stb.

Köszönöm a figyelmet!