

Elosztott intelligens rendszerek
VIMIAC02, 2020. tavasz

SZEMANTIKUS WEB

Méréstechnika és Információs Rendszerek Tanszék

A web napjainkban

- Résztvevők dokumentumokat, adatokat publikálnak, URL címeket adva elérhetővé teszik az információkat
 - Kapcsolódás, hivatkozás linkekkel
- Integrációs próbálkozások: mashup oldalak (ad hoc)
 - Webszolgáltatások integrációja (eltérő API, logika, struktúra)
 - Adatgyűjtés kereső robotokkal (crawler programokkal)
 - Újra és újra felfedezzük a „kerekét”
- Egészítsük ki a webet standard adatelérési módokkal: “Adatweb”

Az adatok webje

- Amire szükségünk van:
 - Publikáljuk adatainkat úgy, hogy felfedezhetőek legyenek a weben:
 - Standardizáljuk az adatok leírását, elérését
 - Dokumentumok eléréhez hasonló, de általános címzés: URI
 - Az URI-kal elérhető forrásokat kapcsoljuk
 - És engedjük a hálózat hatását érvényesülni, ahogy mi is böngészünk a weben...

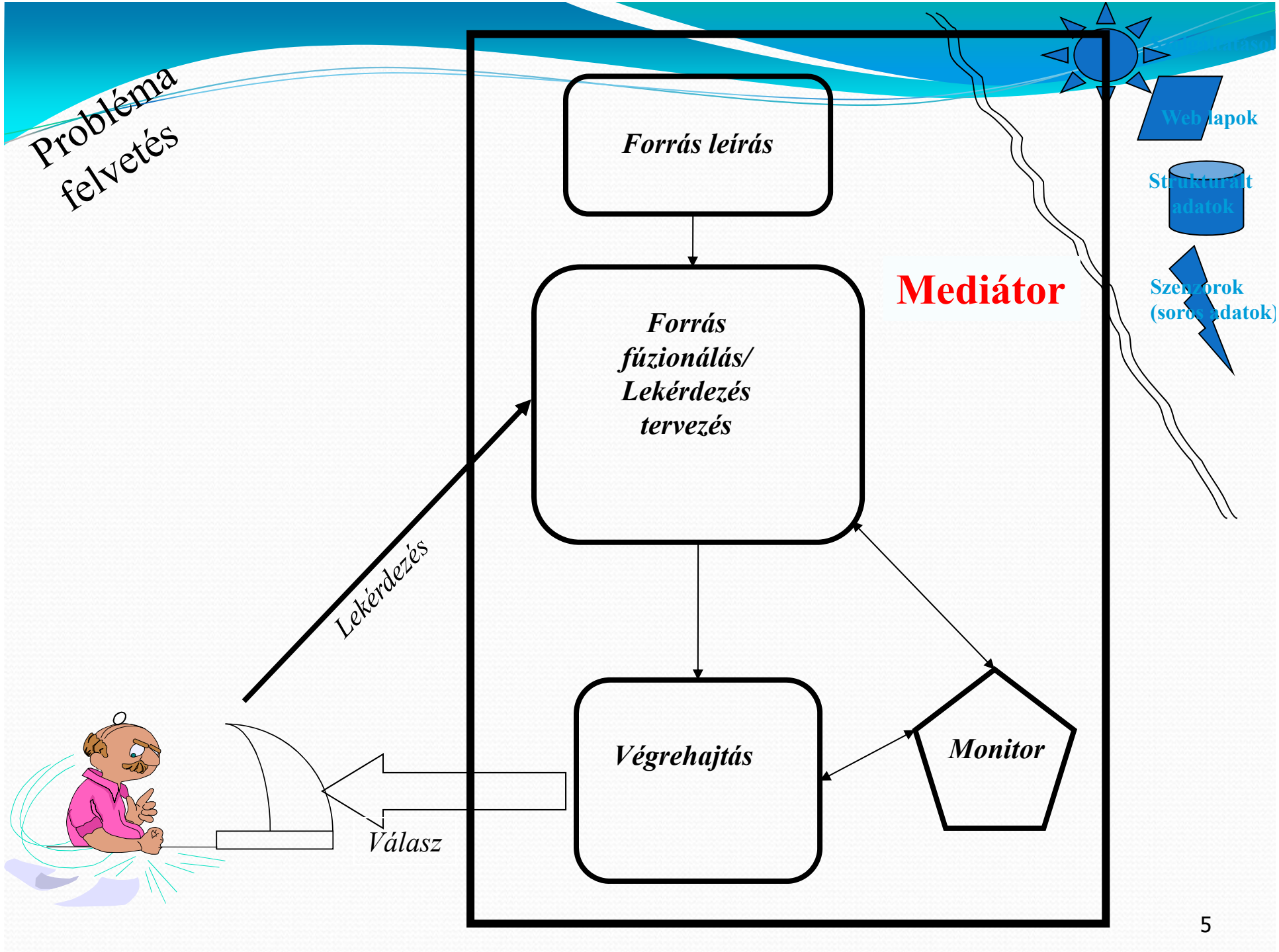
- Példák: 2009 Semantic Technology Conference, San Jose, California, USA, June 15, 2009, Ivan Herman, W3C, ivan@w3.org

Weblapok értelmezése

- A felhasználók megértik, hogy a link egy munkahely honlapjára mutat
- Tudjuk értelmezni, hogy ez egy kutatóhely leírása
- *Ami hiányzik az „adatok webjének” építéséhez:*
 - *jelentéssel kiegészített, értelmezhető linkek*

Tehát bővítsünk:

- Adjunk a linkekhez kiegészítő információt, címkézzük fel őket
- A címkék legyenek géppel értelmezhetőek
 - Kategorizálás
 - Esetleg következtetés



Adatok hálója

- Az adatok webje:
 - Használjunk URI-kat adatok és (nemcsak) dokumentumok publikálására
 - Kapcsoljuk össze az adatokat
 - Jellemezzük/osztályozzuk a linkeket információk hozzáadásával
 - Használjunk standard technológiákat
- *Ez a szemantikus web alapja*



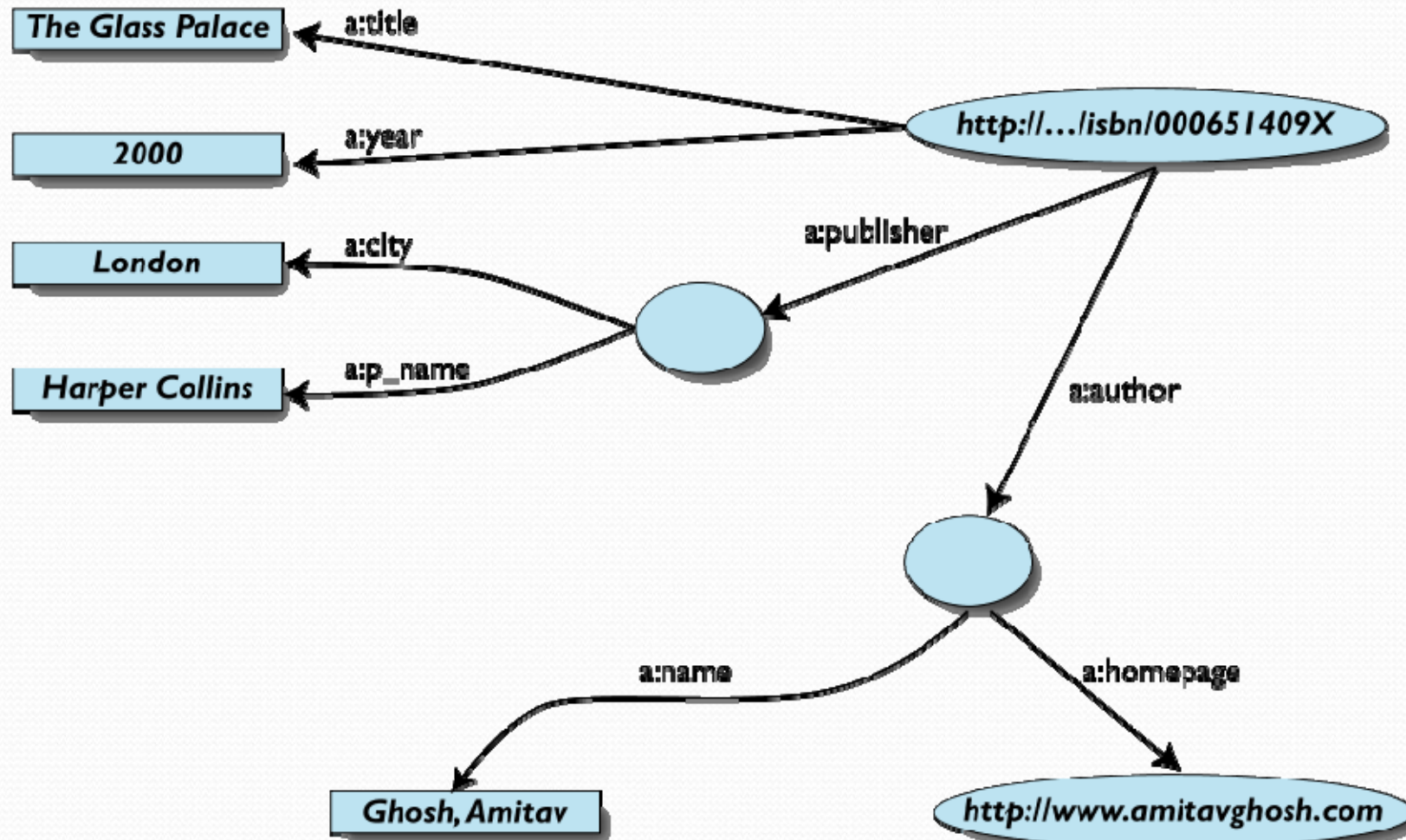
Példa: könyvesbolt adatai

ID	Author	Title	Publisher	Year
ISBN0-00-651409-X	id_xyz	The Glass Palace	id_qpr	2000

ID	Name	Home Page
id_xyz	Ghosh, Amitav	http://www.amitavghosh.com

ID	Publ. Name	City
id_qpr	Harper Collins	London

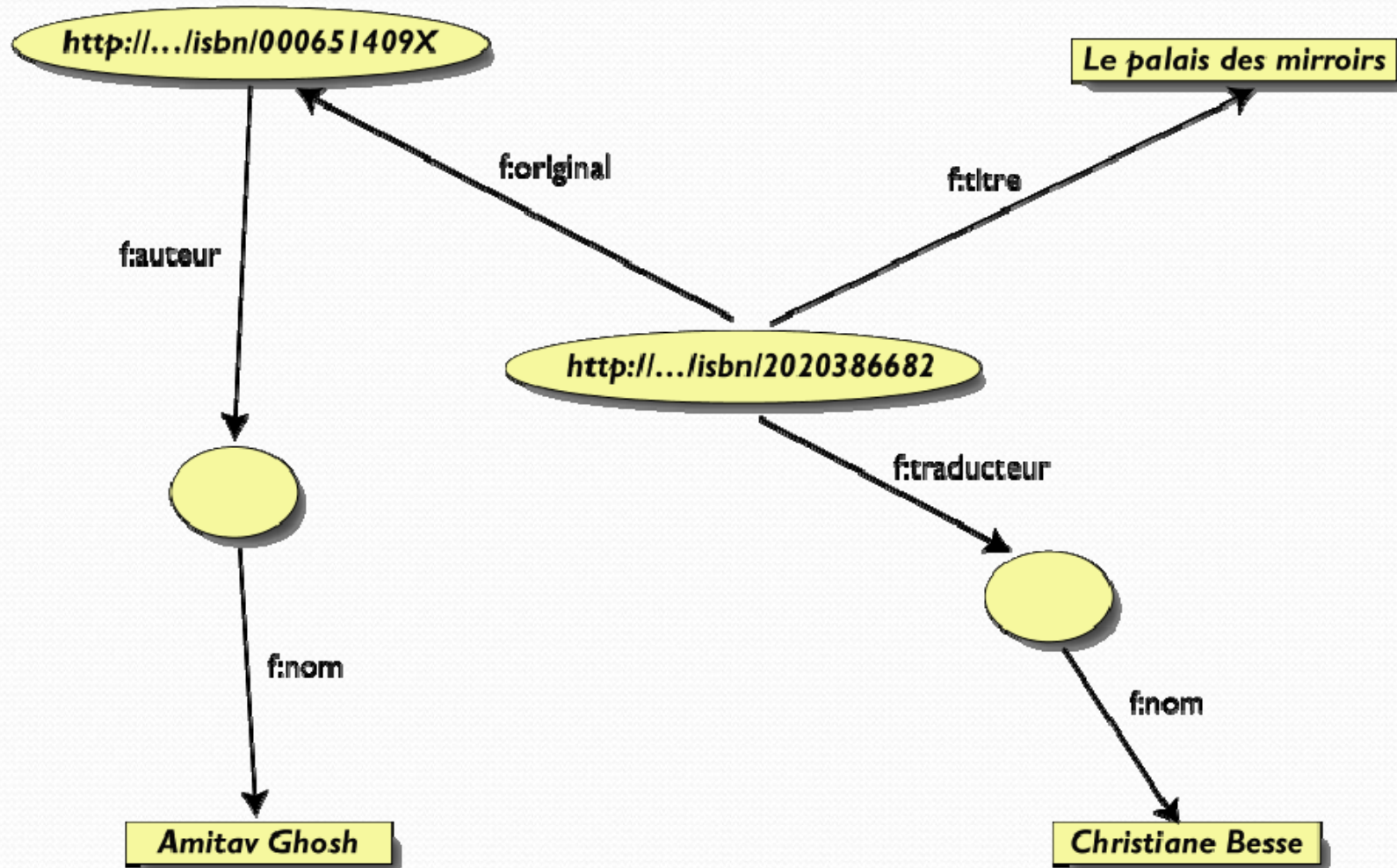
Adatok exportálása relációkként



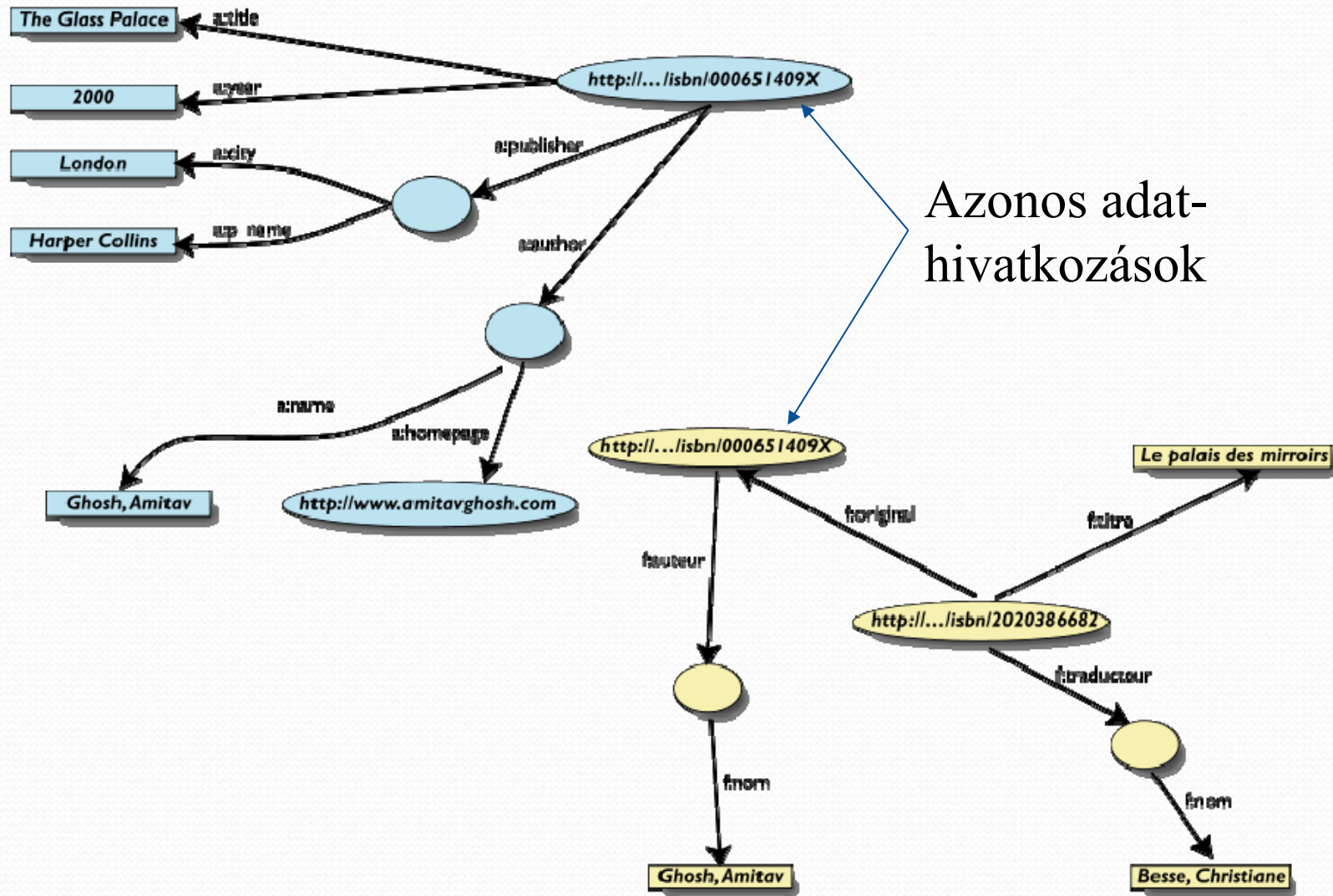
Egy másik könyvesbolt adatai

	A	B	D	E
1	ID	Titre	Traducteur	Original
2	ISBN0 2020386682	Le Palais des miroirs	A13	ISBN-0-00-651409-X
3				
6	ID	Auteur		
7	ISBN-0-00-651409-X	A12		
11	Nom			
12	Ghosh, Amitav			
13	Besse, Christianne			

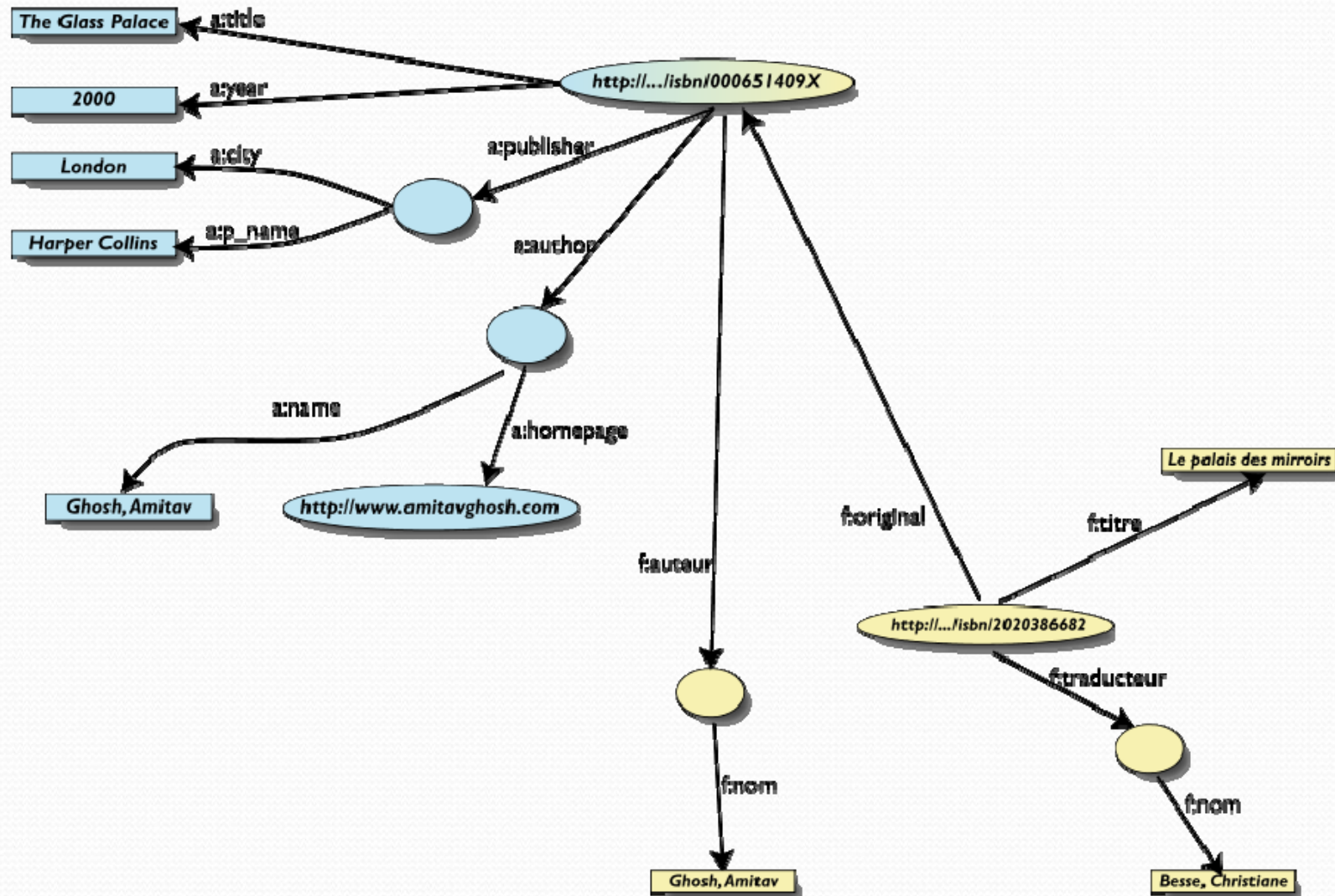
A második könyvesbolt adatainak exportja



Kapcsoljuk össze az adatokat

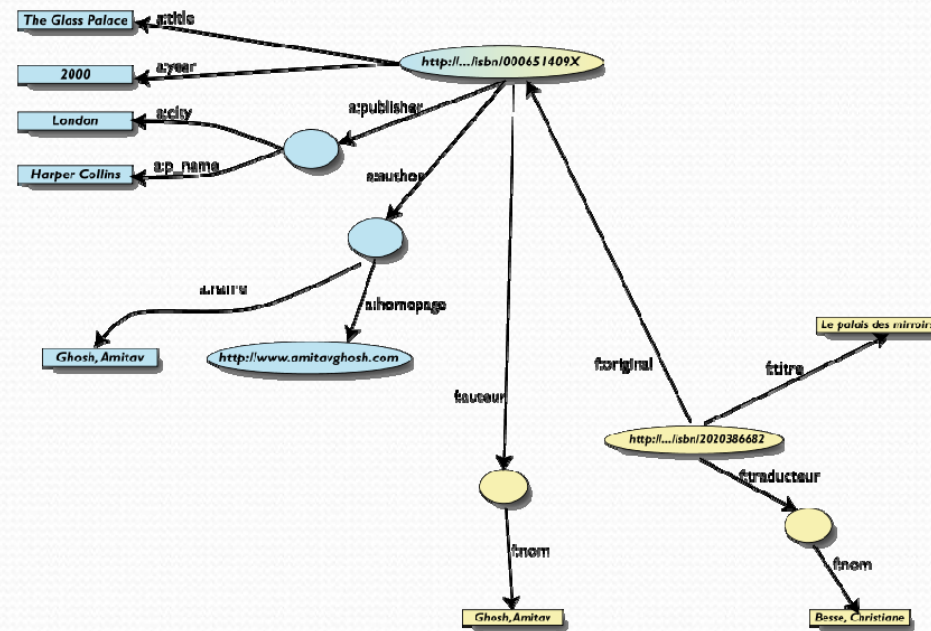


Amennyiben identikusak az elemek:



Írjunk lekérdezéseket a kapcsolatokon keresztül:

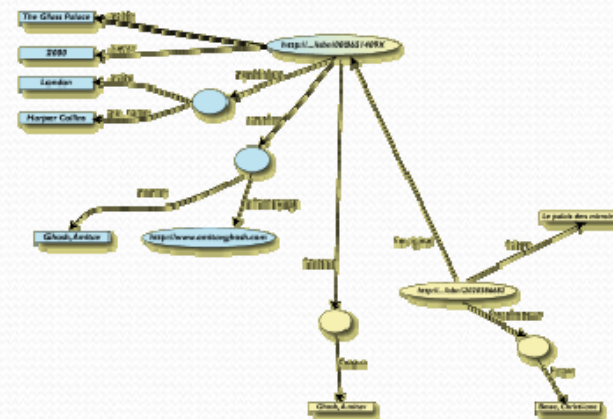
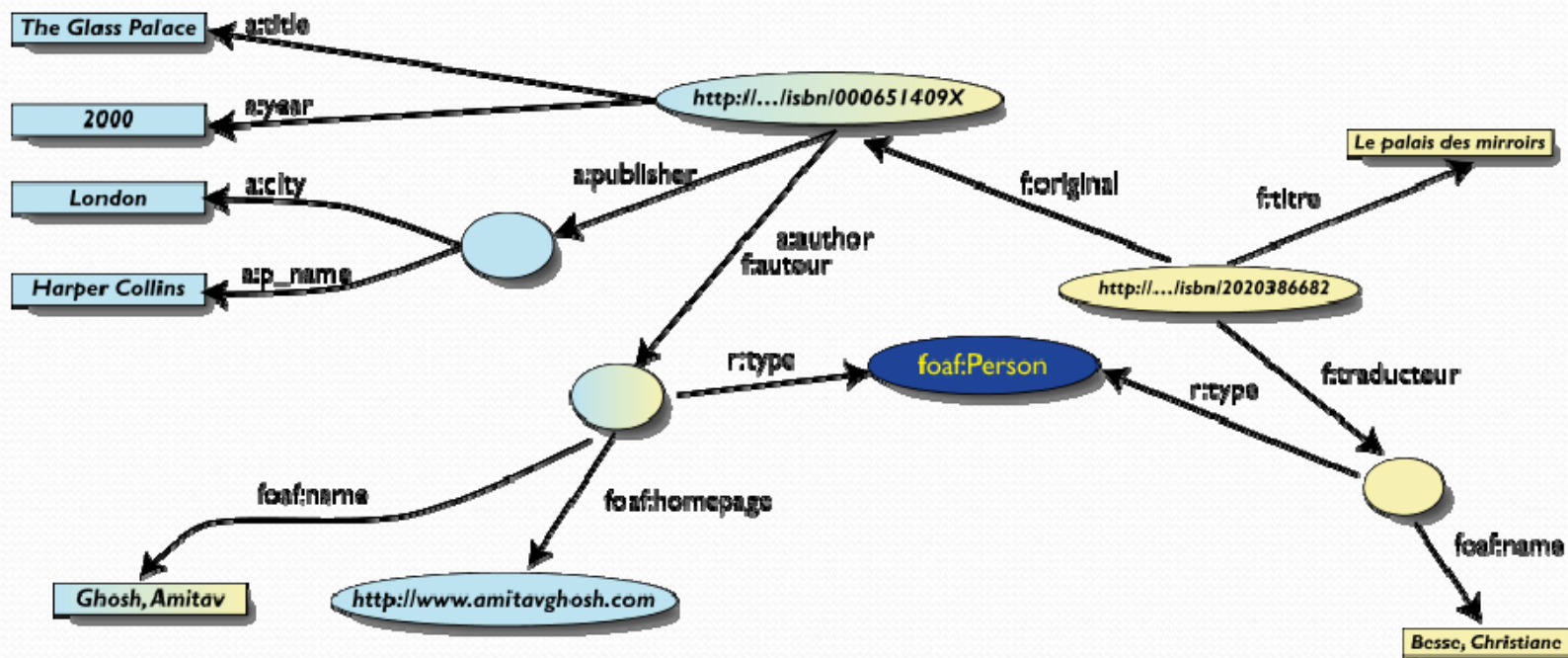
- Az első könyvesbolt adatai kiegészíthető például az eredeti könyvre vonatkozó információkkal



További kapcsolatok is felfedezhetőek...

- Vélhetően az **a : author** és az **f : auteur** azonos elemre mutat
- Automatikus összekapcsoláshoz: adjunk további információt a leíráshoz
 - **a : author** legyen azonos **f : auteur** erőforrással
 - Mindkettő személyt azonosít
 - Ilyen fogalmakat már a webes közösség definiált:
 - egy “Person” elem azonosítható a nevével és a honlapjával
 - Használjuk ezt kategóriaként

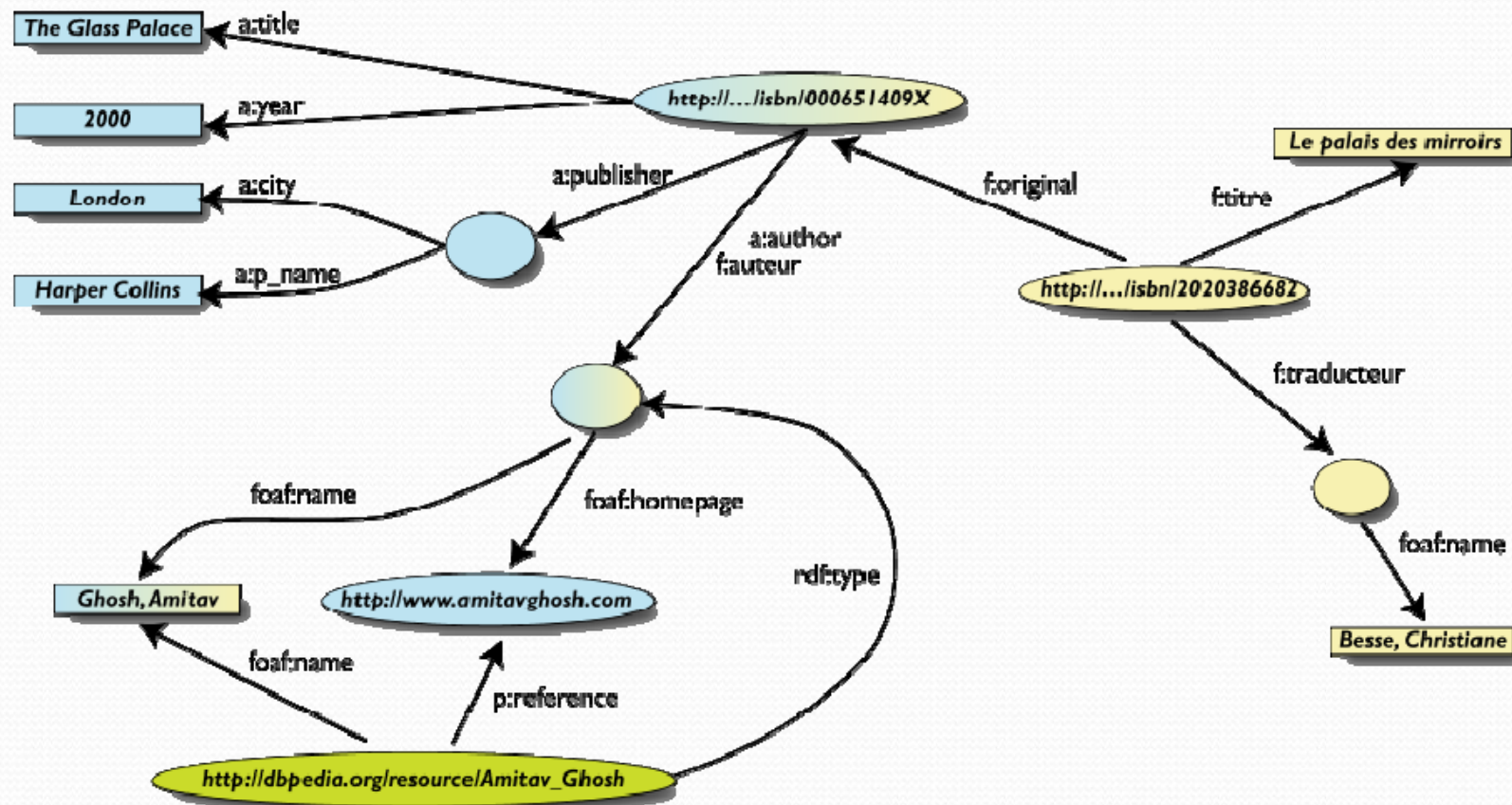
Adatháló kiegészíthető, lekérdezhető így:



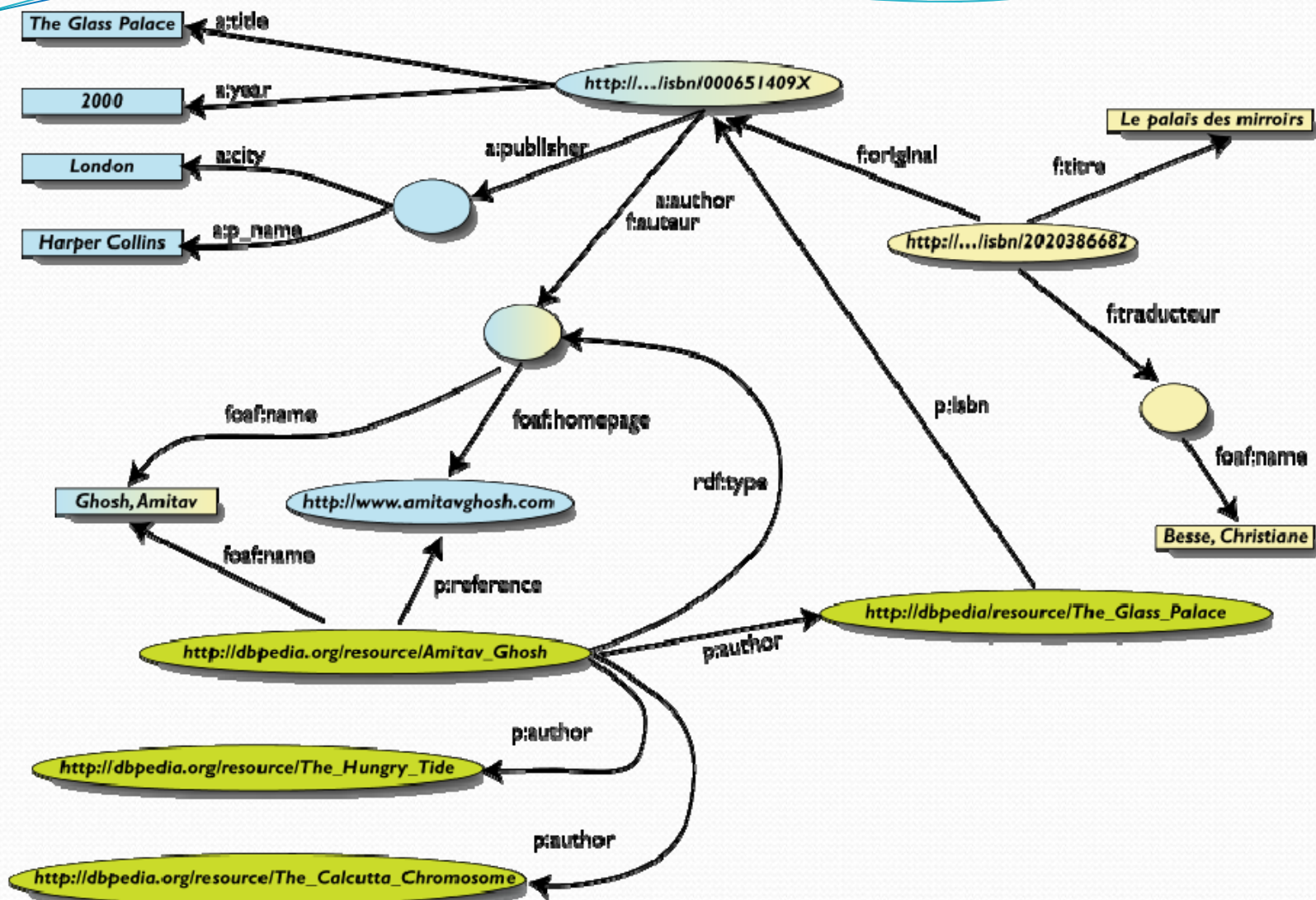
Bővítés újabb adathalmazokkal

- Például a “Person” típus esetén, használhatóak a Wikipédia adatai:
 - pl., a “[dbpedia](#)” projekt már feldolgozta a Wikipédia adatait

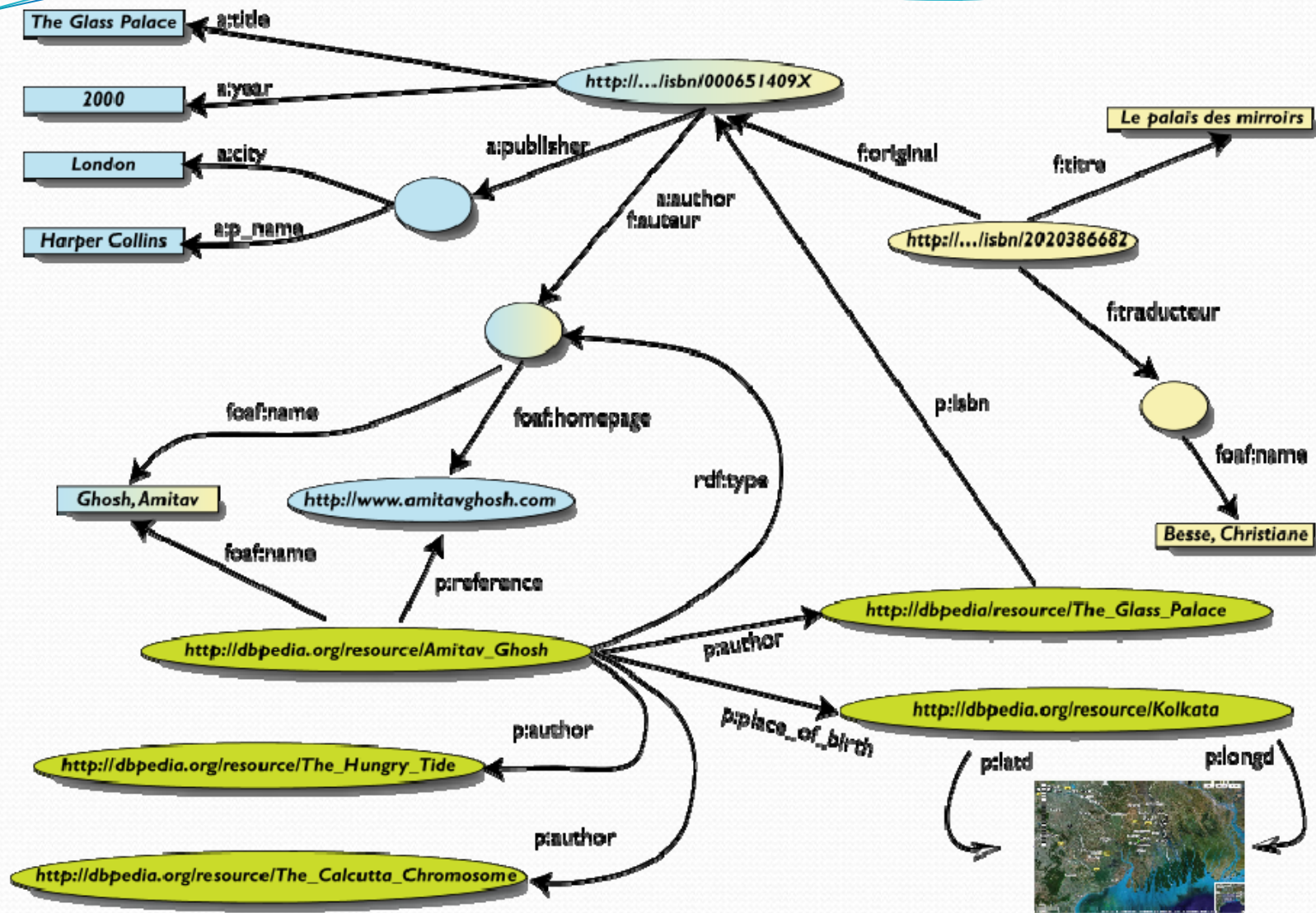
Összekapcsolás Wikipedia adatokkal



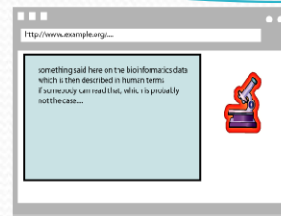
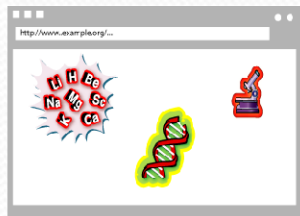
Összekapcsolás Wikipedia adatokkal



És így tovább...

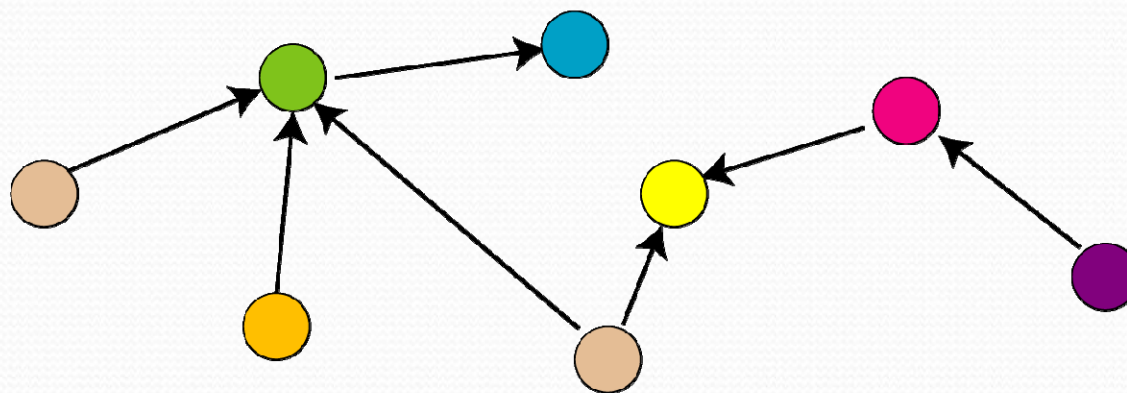


Az adatok webje kialakulófélben...



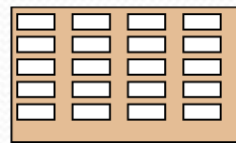
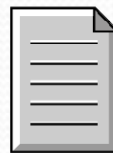
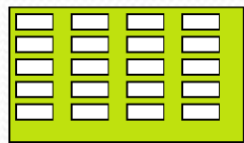
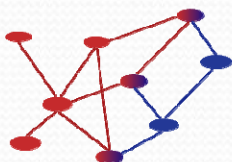
Alkalmazások

Lekérdezés,
adatmódosítás



Absztrakt adatstruktúra

Leképezés,



Forrás adatok különböző formátumokban

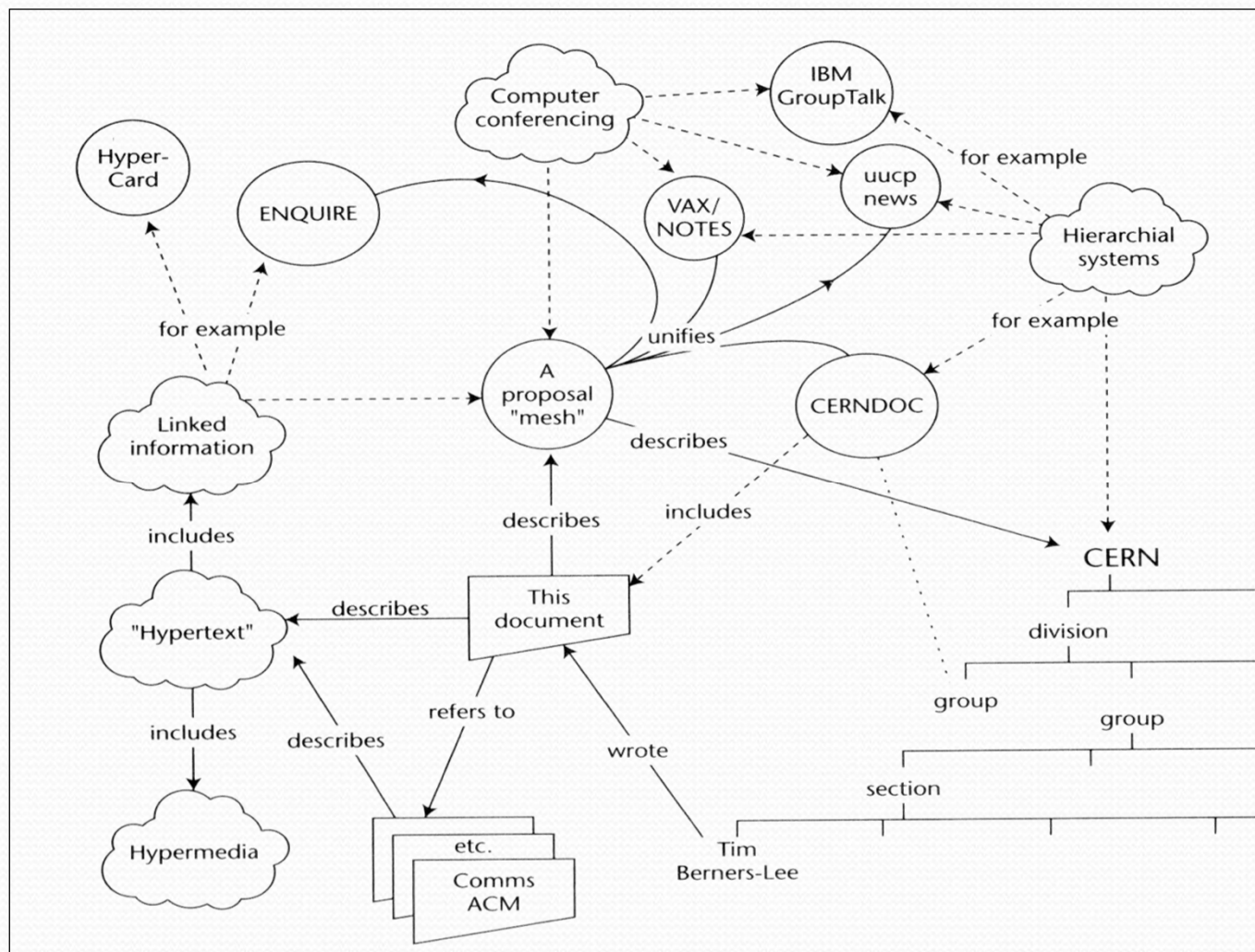
Szemantikus web

Szemantikus Web

- A szemantikus web alkalmas megközelítés, megfelelő nyelvekkel, eszközökkel támogatja az intelligens információs rendszerek fejlesztését az elosztott információs környezetben.
- A SzW alapja a hagyományos web hálózat, így egyáltalán nem nyilvánvaló, hogy alkalmas a feladatra.
- A SzW technológia lehetőséget teremt az ágens alapú intelligens megoldások felhasználására a web területen.

A Szemantikus Web eredete

- Tim Berners-Lee eredeti 1989-es WWW javaslata a Web-et információ menedzselő funkciókkal ellátott objektumok kapcsolataiként jellemzi.



<http://www.w3.org/History/1989/proposal.html>

W3C szervezet célkitűzései

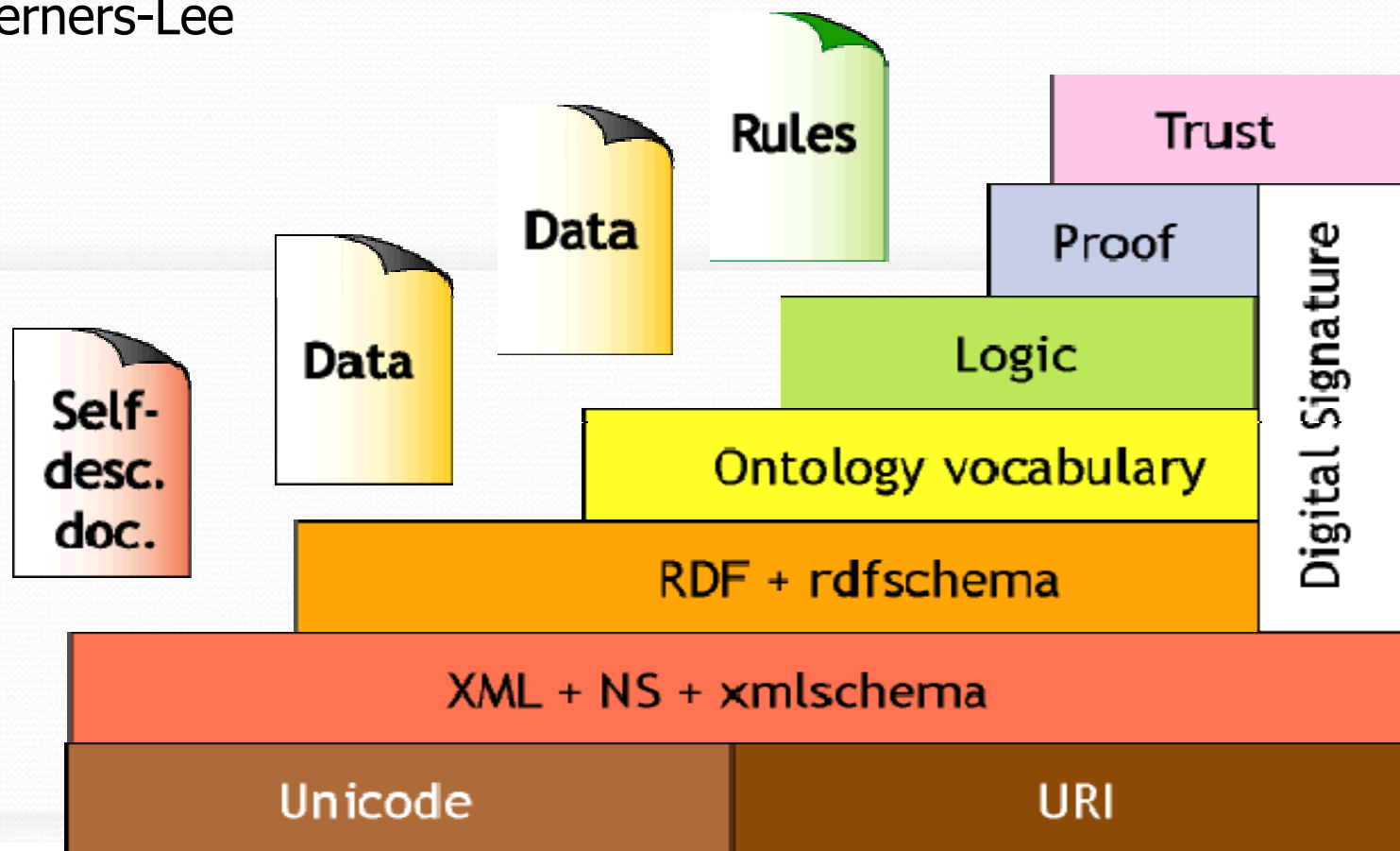
- Megközelítés – számítógépek jobb kihasználtságának biztosítása:

„A szemantikus web egy kiterjesztése a jelenlegi web-nek, amelyben az információknak jól definiált jelentést adhatunk, lehetővé téve a gépek és felhasználók jobb együttműködését..” -- Berners-Lee, Hendler and Lassila, The Semantic Web, Scientific American, 2001

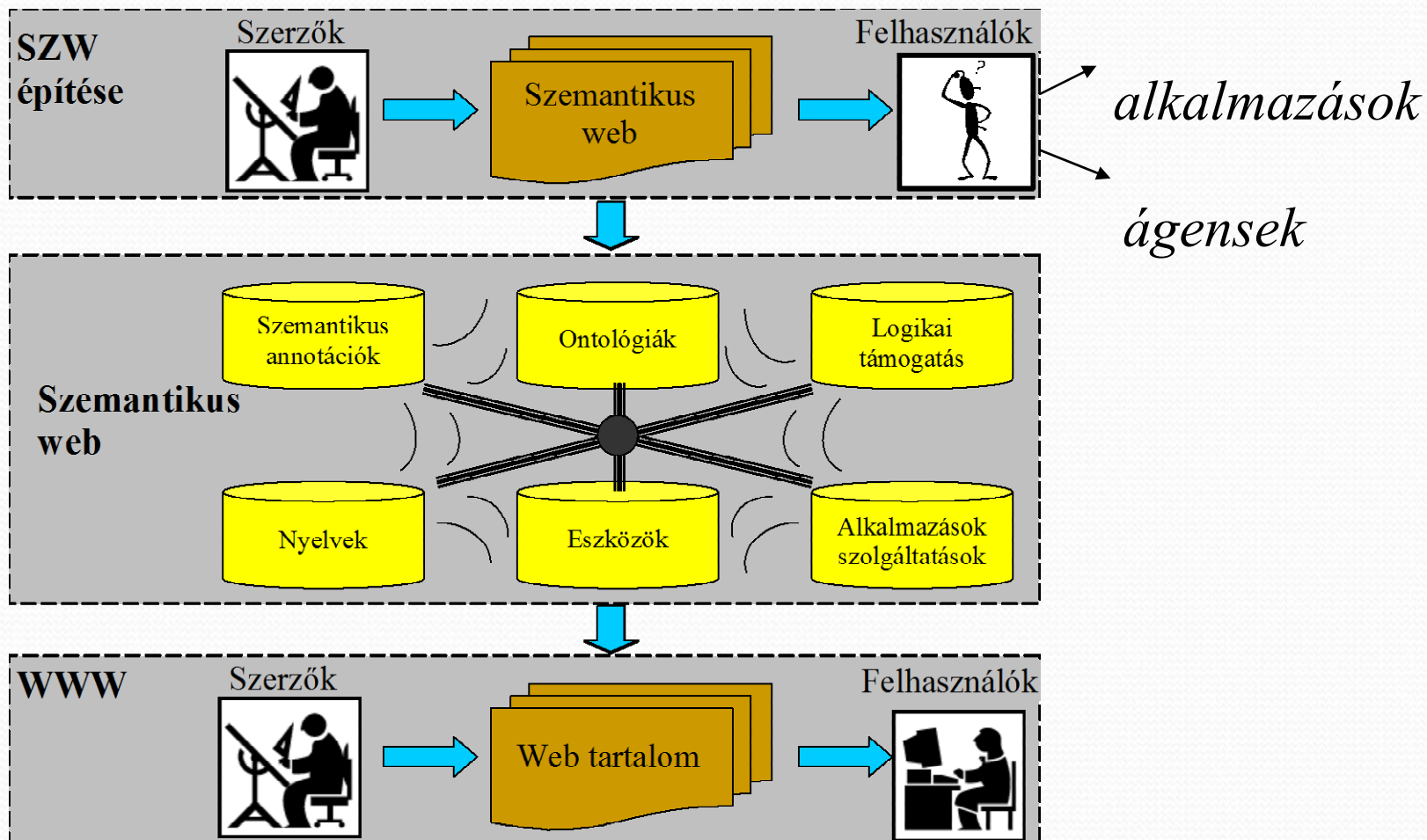
- A jelenlegi web *tárol* dolgokat, míg a szemantikus web képes *működtetni* dolgokat.

TBL szemantikus web felépítése

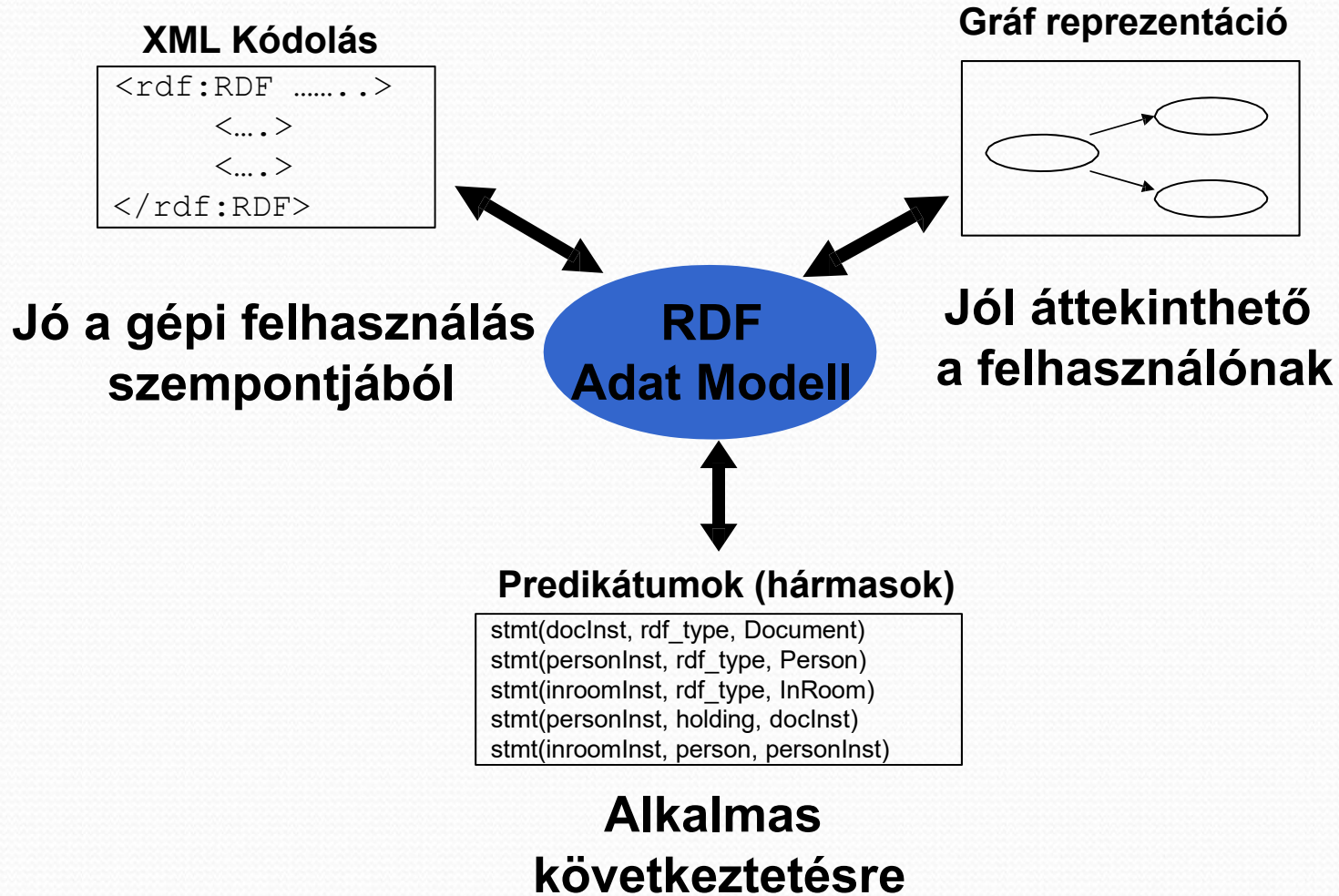
"A szemantikus web elérhetővé teszi a tudást, mint a web a hipertext-et --
Tim Berners-Lee



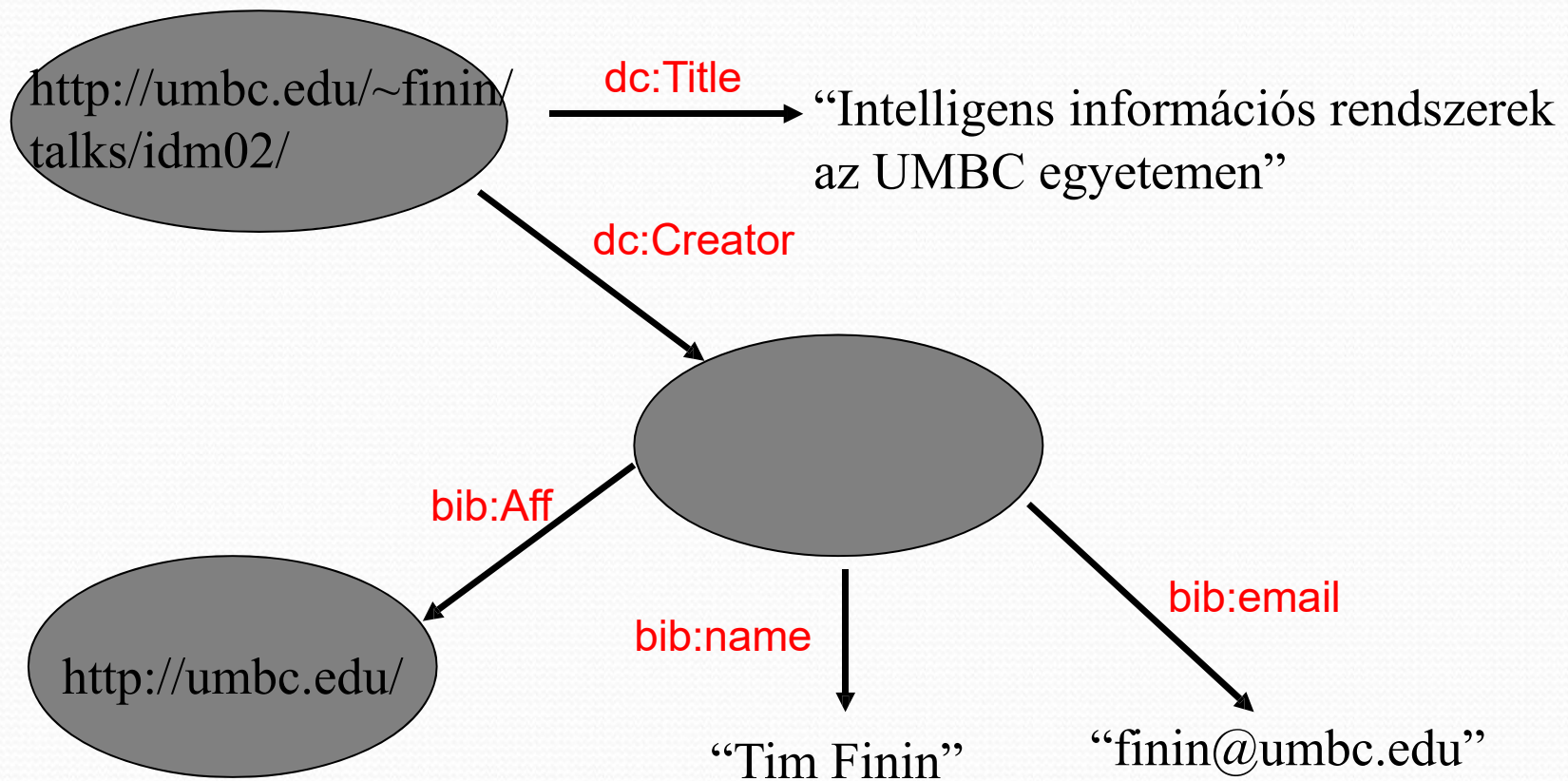
Szemantikus web alkalmazása



RDF - az első SzW nyelv



Egyszerű RDF példa



A példa XML szintaxissal

```
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:bib="http://daml.umbc.edu/ontologies/bib/">
  <description about="http://umbc.edu/~finin/talks/idm02/">
    <dc:title>Intelligent Information Systems on the Web and in the
Aether</dc>Title>
    <dc:creator>
      <description>
        <bib:Name>Tim Finin</bib:Name>
        <bib:Email>finin@umbc.edu</bib:Email>
        <bib:Aff resource=http://umbc.edu/>
      </description>
    </dc:Creator>
  </description>
</rdf:RDF>
```


Hármasokat alkalmazó reprezentáció

- RDF kifejezések leírhatóak
hármasokkal:
- <alany> <állítmány> <tárgy>

Megengedett szintaxis:
<URI><URI><URI>
<URI><URI><string>

RDF tervezési szempontok

- Egyszerű adatmodell
- Formális szemantika és egyszerű következtetési lehetőségek
- Bővíthető URI
- XML alapú szintaktika (is)
 - XML séma adattípusok
- Bárki megfogalmazhat állításokat az erőforrásokról

Alapelvek (1/2)

- Külön értelmezhetően definiálva
 - Modell struktúra (RDF gráf)
 - Interpretációs szemantika (vonzatok)
 - Szintaktikák (XML, TN, N3, ...)
- Mindössze két alap adattípus
 - URI/URIref: minden URI-val azonosított
 - Literálisok
 - String vagy más XSD adattípus

Alapelvek (2/2)

- Integrálható a webes információkkal
 - XML séma adattípusok
 - Referenciák http elérésű információkhoz
- Nyílt világ feltételezés
 - Bárki megfogalmazhat állításokat bármilyen erőforráshoz
 - Nem garantált a teljesség
 - Nem garantált a konzisztencia

Alapelemek

- Gráf adatmodell
- URI alapú szótárak
- Adattípusok
- Literálisok
- XML szerializációs szintaktika
- Egyszerű tények leírása
- Következtetés

Linked Open Data projekt

- Cél: elérhetővé tenni a nyitott RDF adatbázisokat
- *Hozzunk létre kapcsolatokat az RDF adathalmazok között*
- Rögzítsünk lekérdezési lehetőségeket
- Ma is elérhetőek: milliárdnyi hármások, sok millió kapcsolat



Linked Open Data projekt

- A szemantikus web alapja a nagy mennyiségű tartalom szerint kapcsolható, a webes elérhető RDF adatok létezése
- A Linked Data nem specifikáció, hanem jó gyakorlatok gyűjteménye adatok megosztására a weben
- További szemantikus web technológiák (RDFS, OWL, SPARQL) segítik az adatok alkalmazásokba építését

5-star data



5-star data

Adatok nyílt licenz mentén elérhetőek a weben, de dokumentumokba „zárva”



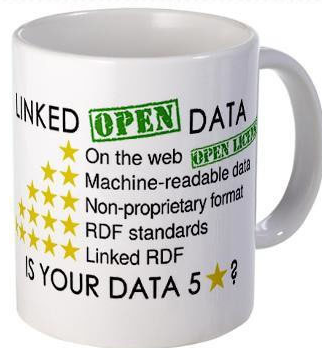
5-star data

Adatok nyíltan elérhetőek,
itt már struktúrált formában,
de „zárt”, licensszel védett
technológiákon keresztül



5-star data

Adatok nyíltan elérhetőek,
nyílt formátumban,
de még mindig csak
egyszerű adatok a weben
és nem az adatok
webje...

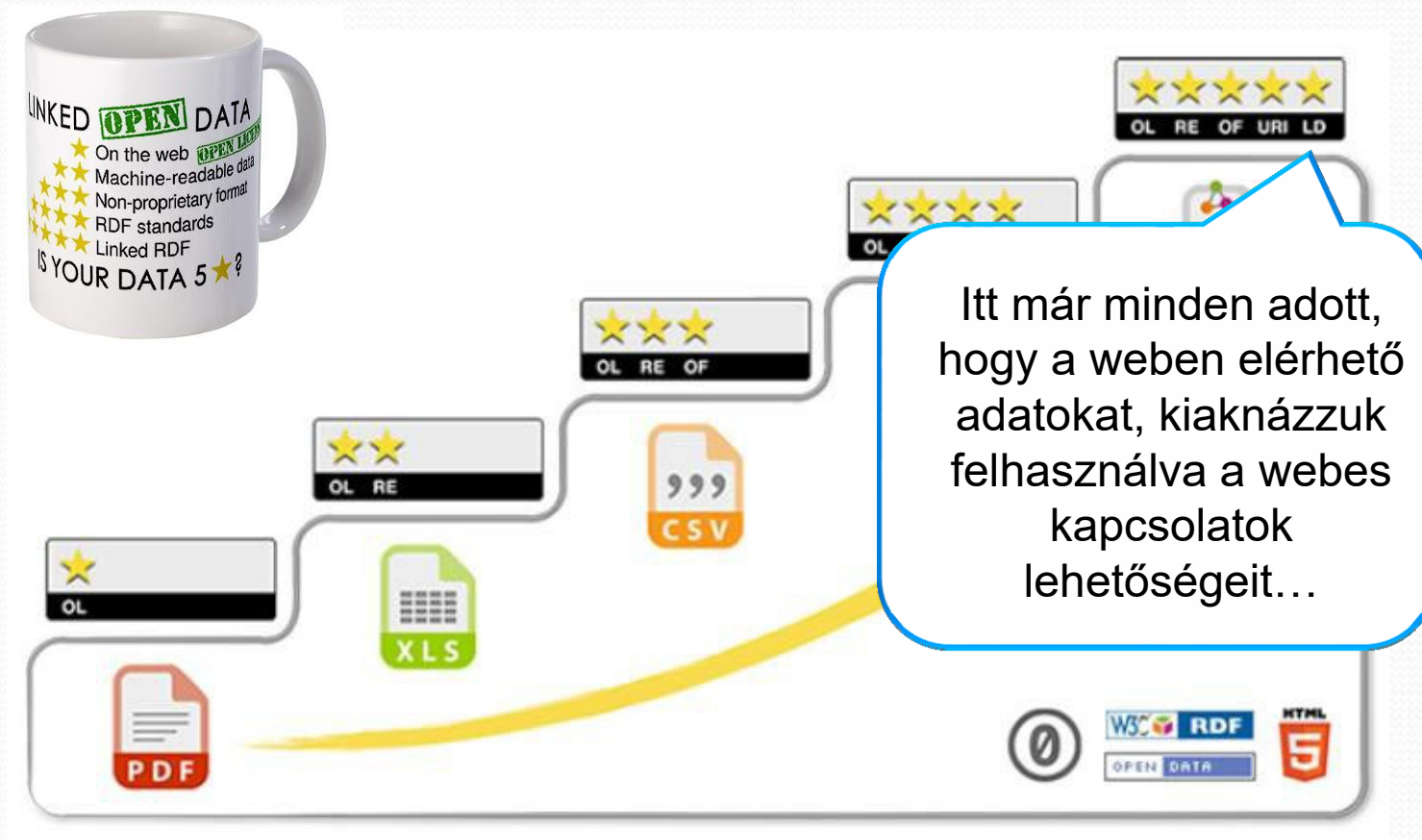


5-star data

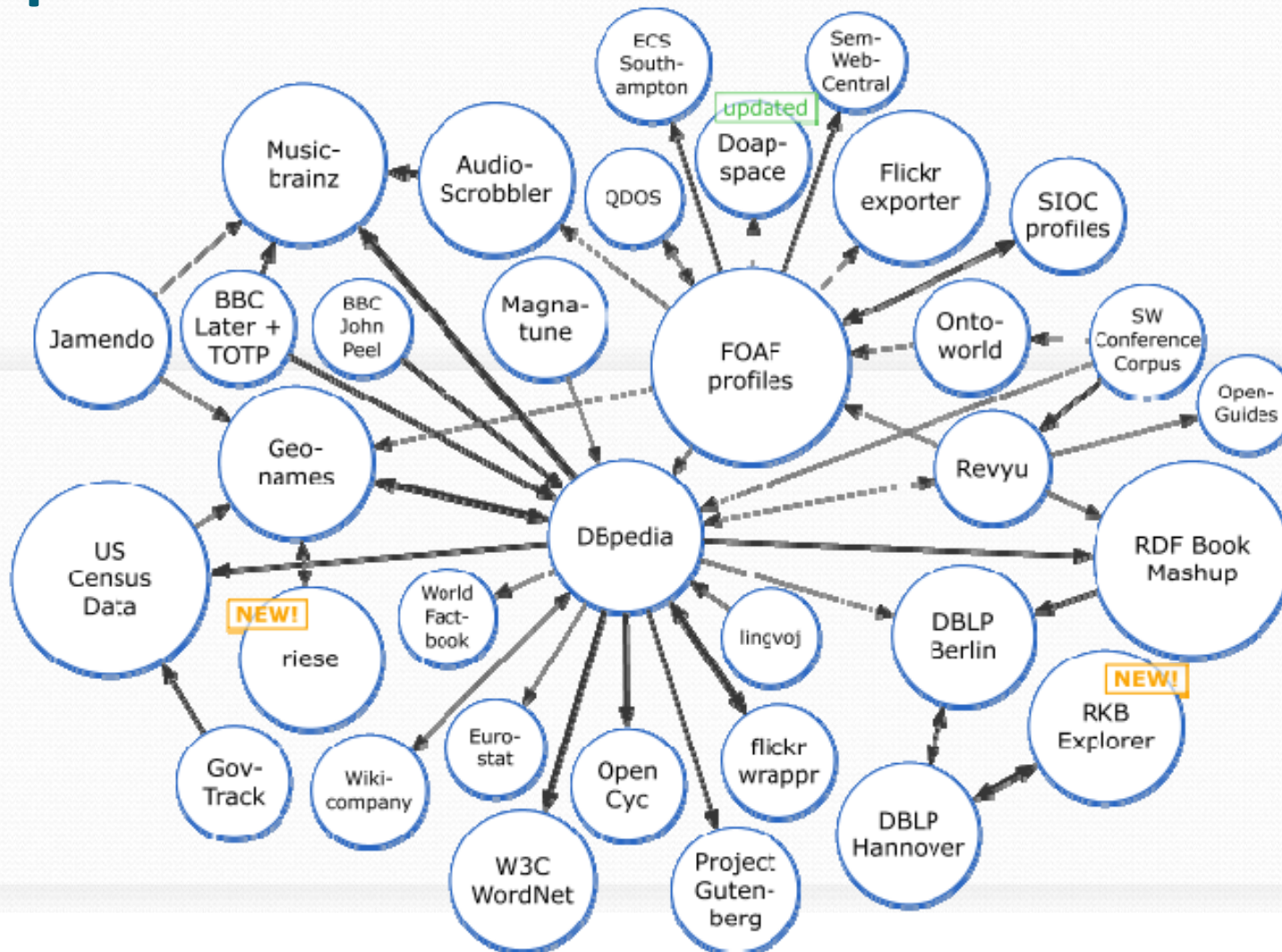
Itt már az adatok közvetlenül elérhetőek a weben az egyes forrásokban, de hiányzanak a kapcsolatok



5-star data



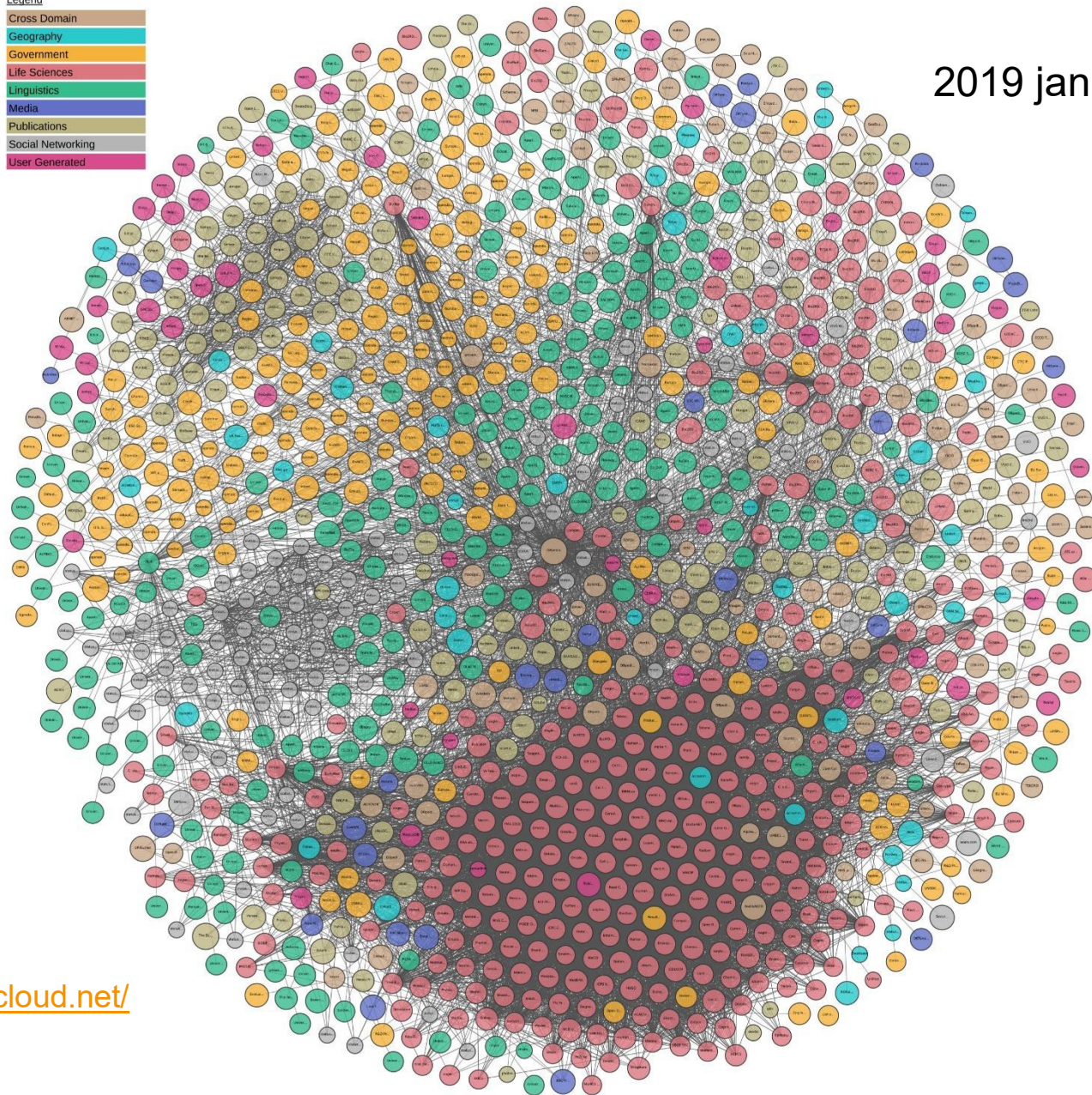
A kapcsolati felhő – kiindulás 2008



2019 január

Legend

Cross Domain
Geography
Government
Life Sciences
Linguistics
Media
Publications
Social Networking
User Generated



<https://lod-cloud.net/>

The Linked Open Data Cloud from lod-cloud.net

1/15/2019



Tehát mi is a Linked Data?

- A Linked Data megközelítés (szabvány) realizálja az eredeti Szemantikus Web koncepciót az összekötöttség tekintetében
- Összekapcsol információ forrásokat FOAF, RDF, OWL és egyéb formátumokban

A Linked Data szabályai

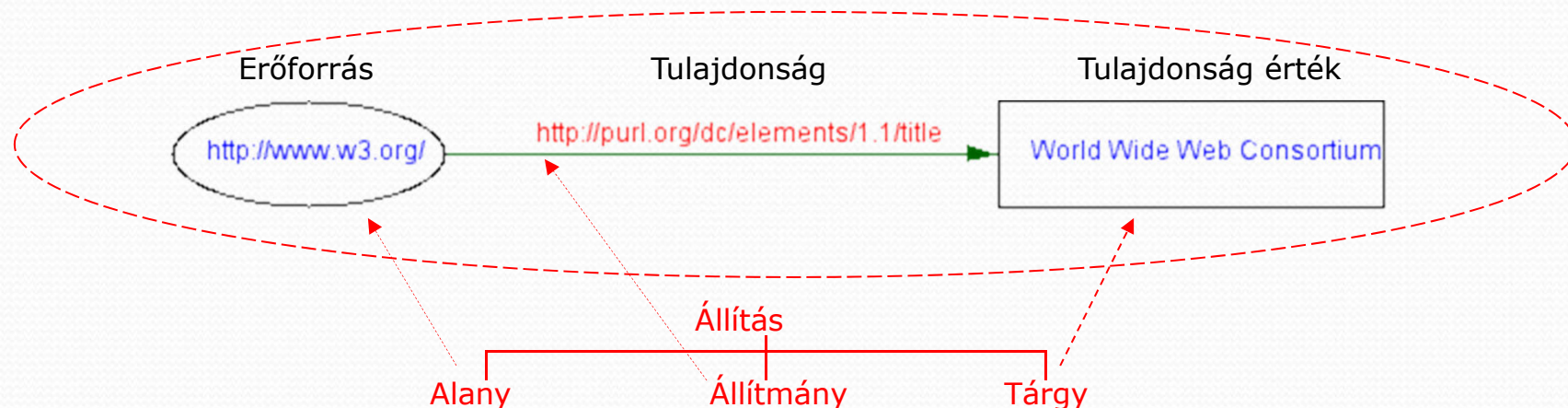
- Minden adat megnevezése URI- keresztül
- A kapcsolatok URI-jai legyenek érvényes URL-ek
- Legyen egy URL lap, amely tartalmazza az URI-val hivatkozott adatot, ennek a lapnak az URL címe ne változzon
- Az adatok dokumentumokban, fájlokban legyenek elérhetőek a weben

Célok

- Számos forrás tartalmaz hasonló adatokat a weben
 - Pl. személyes adatok
- Az adatokat kössük össze fogalomtárakon keresztül, számos fogalomtár egyidejű felhasználásával.
- Az eredmény egy hatalmas háló, gráf.

RDF

- Eredetileg webes metaadat kezelés javasolt megközelítése volt.
- Objektum típusok: erőforrás, tulajdonság, és állítások.
- RDF XML szintaxis
- Alkalmazási terület független

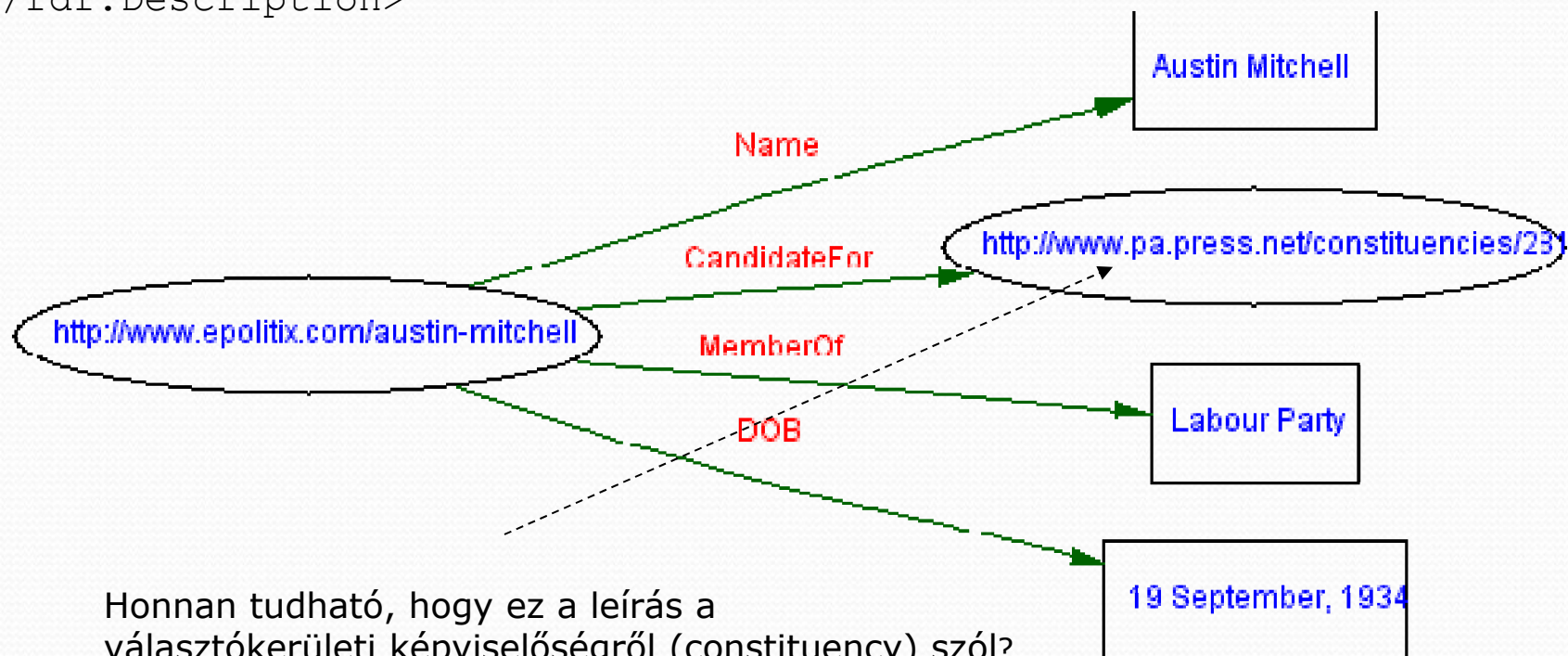


Mire használható az RDFS

- Tekinthető RDF kifejezés szótárként is, hagyományos megközelítésként séma
- Alkalmazási területek szemantikájának leírása.
- Tulajdonságok újrafelhasználhatóságának biztosítása. Értelmezési tartomány és értékkészlet megadása
- Tulajdonságok osztályokhoz, alosztályokhoz rendelése specifikálható

Adat integrációs probléma

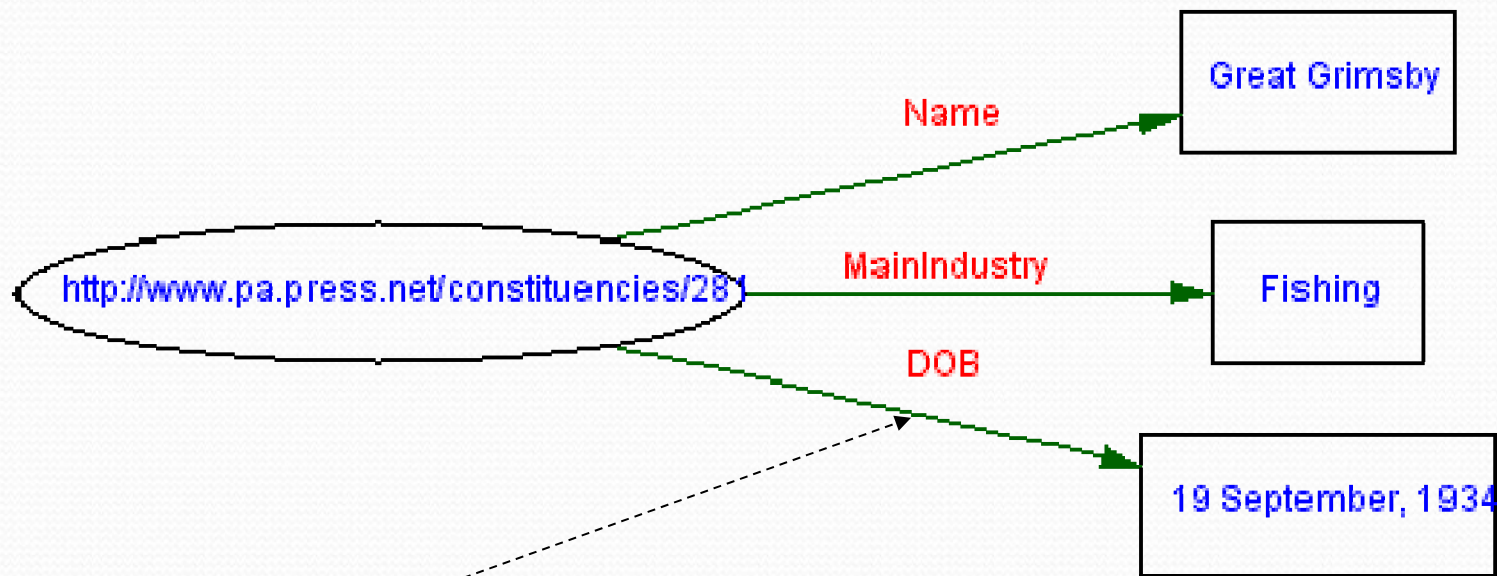
```
<rdf:Description rdf:about="http://www.epolitix.com/austin-mitchell">  
  <epx:Name>Austin Mitchell</epx:Name>  
  <epx:CandidateFor  
    rdf:resource="http://www.pa.press.net/constituencies/281" />  
  <epx:MemberOf>Labour Party</epx:MemberOf>  
  <epx:DOB>19 September, 1934</epx:DOB>  
</rdf:Description>
```



Honnan tudható, hogy ez a leírás a választókerületi képviselőségről (constituency) szól?

Adat integrációs probléma

```
<rdf:Description rdf:about="http://www.pa.press.net/constituencies/281">  
  <epx:Name>Great Grimsby</epx:Name>  
  <epx:MainIndustry>Fishing</epx:MainIndustry>  
  <epx:DOB>19 September, 1934</epx:DOB>  
</rdf:Description>
```



Egy választókerületnek van születési dátuma?

Állítások validációja

- Alkalmas-e az alany az állításhoz?

```
{  
  epx:CandidateFor,  
  [http://www.epolitix.com/austin-mitchell],  
  [http://www.microsoft.com/]  
}
```

- Alkalmas-e az állítmány az alanyhoz?

```
{ epx:DOB,  
  [http://www.pa.press.net/constituencies/281],  
  "19 September, 1934" }
```

Erőforrások típusizálása

- Honnan tudjuk, hogy ez egy választókörizet?

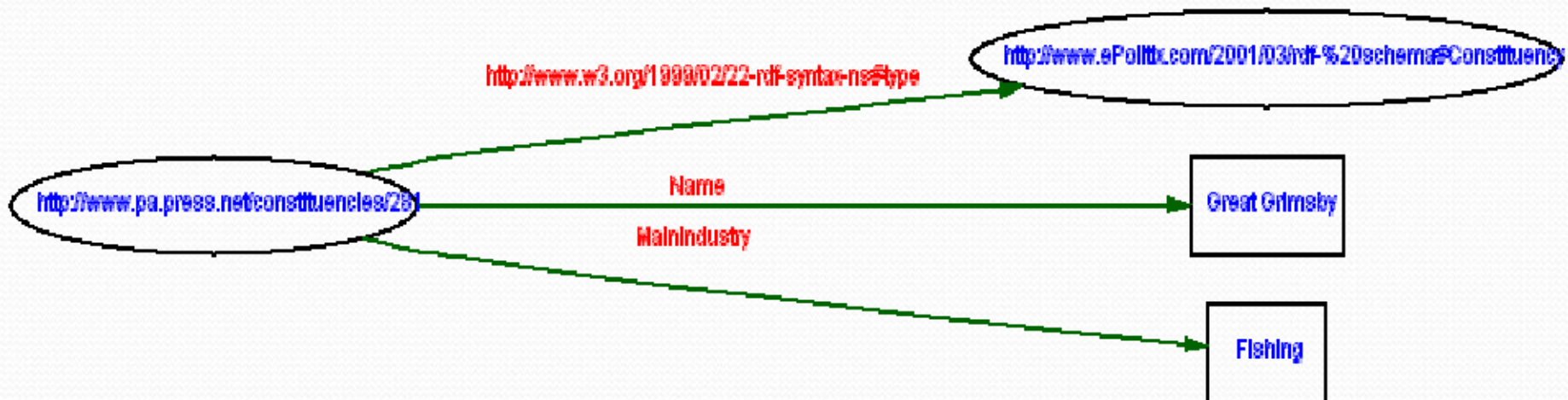
```
<rdf:Description rdf:about="http://www.pa.press.net/constituencies/281">  
  <epx:Name>Great Grimsby</epx:Name>  
  <epx:MainIndustry>Fishing</epx:MainIndustry>  
</rdf:Description>
```

- „választókerület” típusba kell sorolni az erőforrást.

Erőforrás típus: Consistency

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:epx="http://www.ePolitix.com/2001/03/rdf-schema#" >

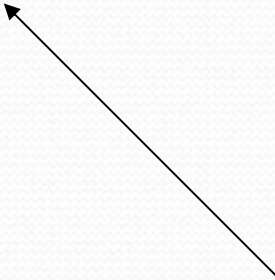
  <rdf:Description rdf:about=
    "http://www.pa.press.net/constituencies/281">
    <rdf:type resource=
      "http://www.ePolitix.com/2001/03/rdf-schema#Constituency"/>
    <epx:Name>Great Grimsby</epx:Name>
    <epx:MainIndustry>Fishing</epx:MainIndustry>
  </rdf:Description>
</rdf:RDF>
```



Tulajdonság értékek specifikálása: megfelelő állítmányok használata az alanyokhoz

```
{  
  epx:CandidateFor,  
  [http://www.epolitix.com/austin-mitchell],  
  [http://www.microsoft.com/]  
}
```

Érvénytelen: Nem választókerület



Tulajdonság értékek specifikálása: megfelelő állítmányok használata az alanyokhoz

Hozzuk létre: *rdf:property CandidateFor*

```
<rdf:Description
  rdf:about="http://www.ePolitix.com/2001/03/rdf-schema#CandidateFor"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#" >
  <rdf:type
    resource="http://www.w3.org/1999/02/22-rdf-syntax-ns#Property" />
</rdf:Description>
```



Tulajdonság értékek specifikálása: megfelelő állítmányok használata az alanyokhoz

ACandidateFor tulajdonság használatát korlátozzuk, hogy csak Constituency típushoz legyen alkalmazható

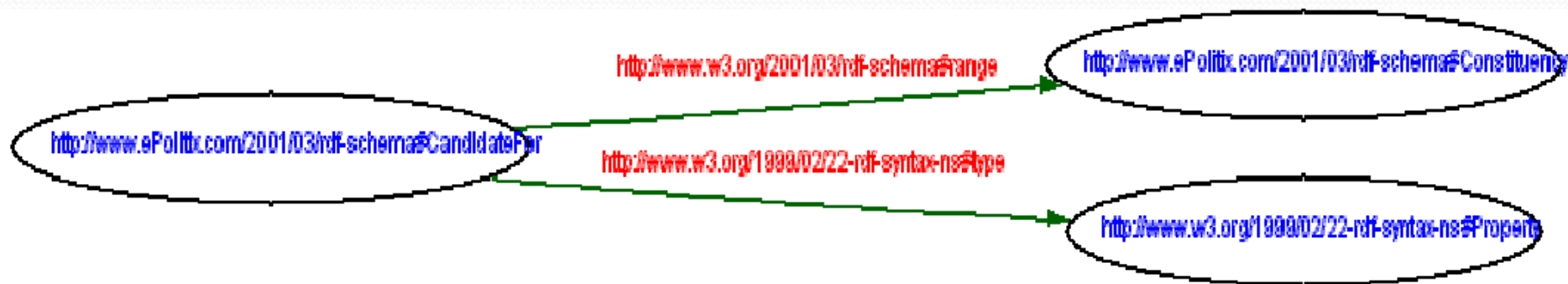
```
<rdf:Property
  rdf:about="http://www.ePolitix.com/2001/03/rdf-schema#CandidateFor" >
  <rdfs:range rdf:resource=
    "http://www.ePolitix.com/2001/03/rdf-schema#Constituency" />
</rdf:Property>
```



RDFS Névterek

Az RDFS séma névteréhez tartozó prefix,
URI: <http://www.w3.org/2000/01/rdf-schema#>

```
<rdf:Property
  rdf:about="http://www.ePolitix.com/2001/03/rdf-schema#CandidateFor">
  <rdfs:range rdf:resource=
    "http://www.ePolitix.com/2001/03/rdf-schema#Constituency"/>
</rdf:Property>
```



rdfs:range

A *ConstraintProperty* használata megadja, hogy az adott osztályok milyen tulajdonságúak lehetnek.

A *range* tulajdonság értéke minden esetben egy osztály. Egy tulajdonsághoz maximum egy *range* értéket lehet megadni.

```
<rdf:Property
  rdf:about="http://www.ePolitix.com/2001/03/rdf-schema#CandidateFor" >
  <rdfs:range rdf:resource=
    "http://www.ePolitix.com/2001/03/rdf-schema#Constituency" />
</rdf:Property>
```



Kifejezések validációja

- Alkalmas-e az alany az állításhoz?

```
{
  epx:CandidateFor,
  [http://www.epolitix.com/austin-mitchell],
  [http://www.microsoft.com/]
}
```

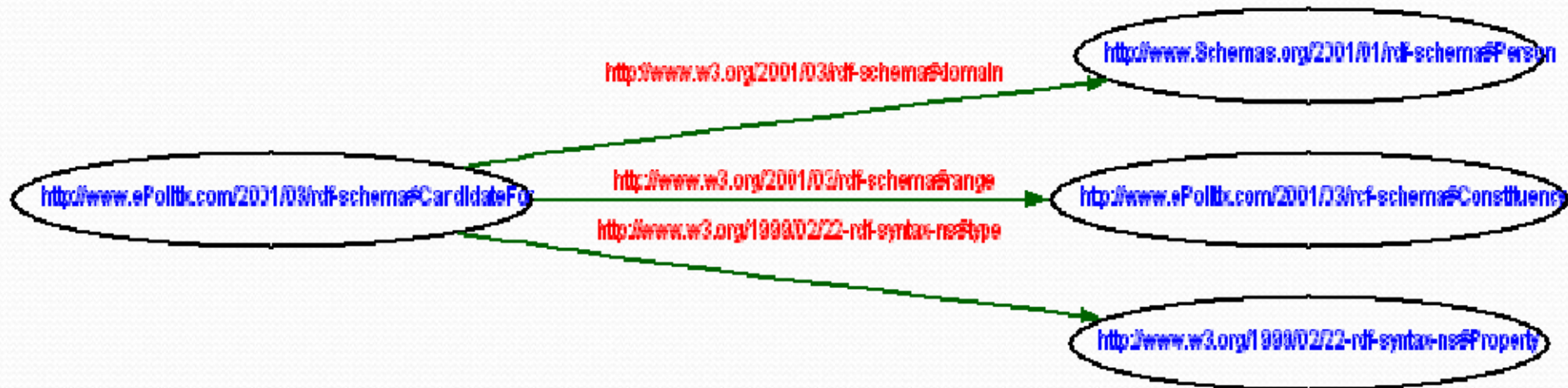
- Az állítmány alkalmazható-e az alanyhoz?

```
{ epx:DOB, [http://www.pa.press.net/constituencies/281],
  "19 September, 1934" }
```

Erőforrás tulajdonságok specifikálása: állítmányok meghatározása alanyokhoz

Korlátozzuk a *CandidateFor* tulajdonságot, hogy csak *Person* típusú alanyokhoz legyen használható

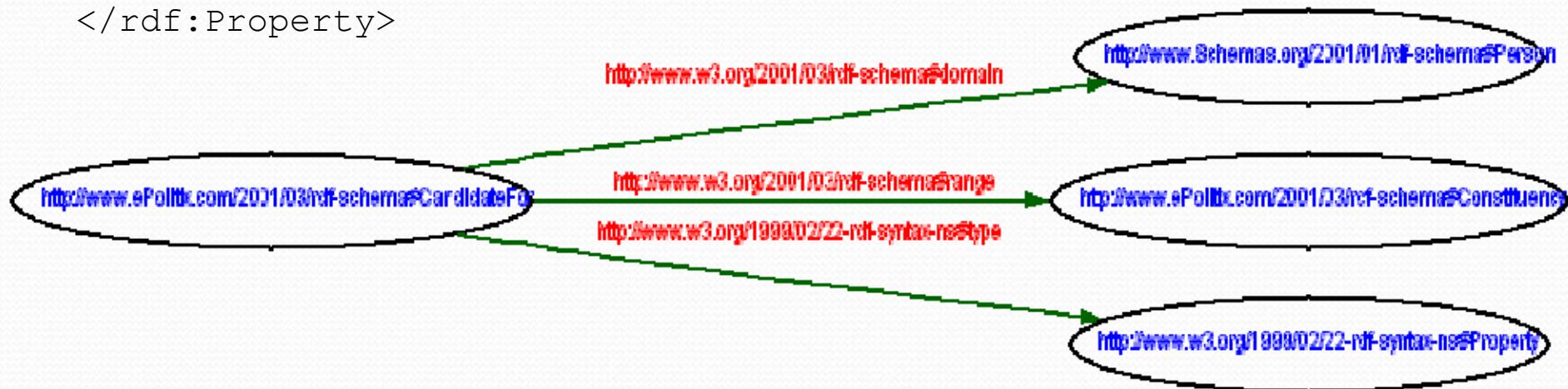
```
rdf:Property
  rdf:about="http://www.ePolitix.com/2001/03/rdf-schema#CandidateFor">
  <rdfs:domain
    rdf:resource="http://www.Schemas.org/2001/01/rdf-schema#Person" />
  <rdfs:range
    rdf:resource=
      "http://www.ePolitix.com/2001/03/rdf-schema#Constituency" />
</rdf:Property>
```



rdfs:domain

A *domain* tulajdonság értéke minden esetben egy osztály. Egy tulajdonsághoz maximum egy *domain* értéket lehet megadni.

```
rdf:Property
  rdf:about="http://www.ePolitix.com/2001/03/rdf-
  schema#CandidateFor">
  <rdfs:domain
    rdf:resource="http://www.Schemas.org/2001/01/rdf-
  schema#Person" />
  <rdfs:range
    rdf:resource="http://www.ePolitix.com/2001/03/rdf-
  schema#Constituency" />
</rdf:Property>
```



Specifikációk:

`rdfs:ConstraintProperty`

Ez az erőforrás az *rdf:Property* alosztályát definiálja, melynek példányai kényszerek megadására alkalmazható tulajdonságok.

Ez az osztály alosztálya az *rdfs:ConstraintResource* osztálynak, ennek az osztálynak azt a részhalmazát definiálja, amelyek tulajdonságok.

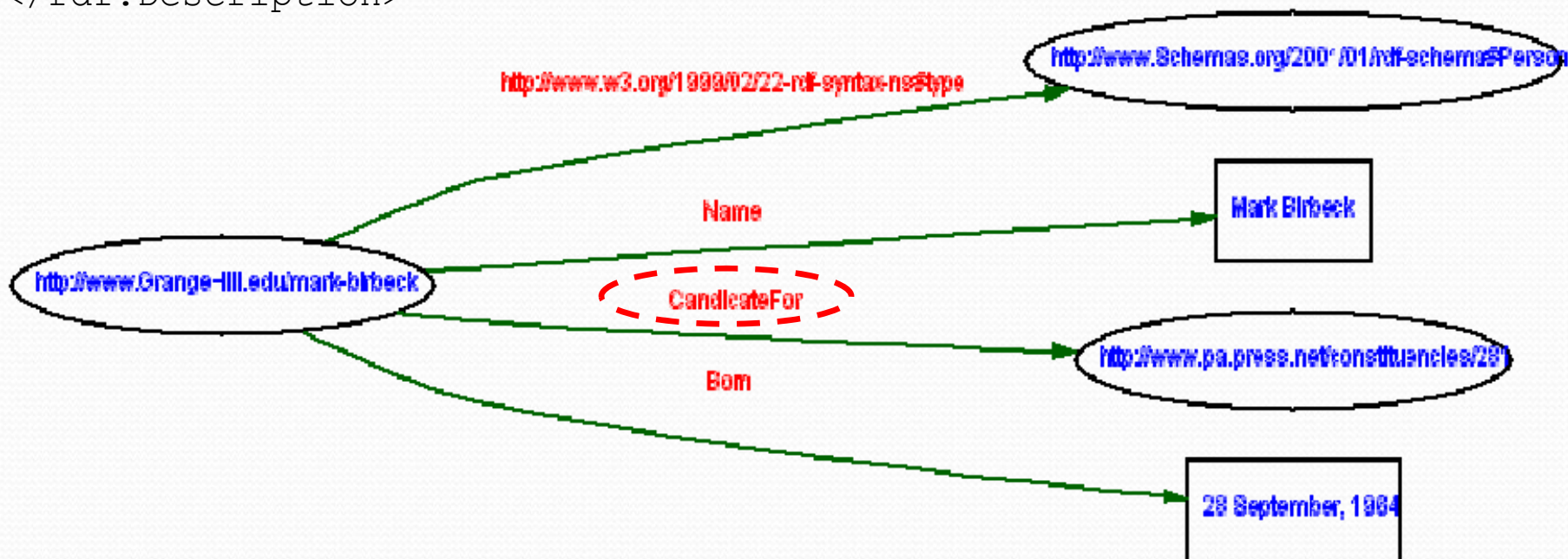
Az *rdfs:domain* és az *rdfs:range* példányai *rdfs:ConstraintProperty* osztálynak.

`rdfs:ConstraintResource`

Ez az erőforrás az *rdfs:Resource* alosztályát definiálja, melynek példányai olyan RDFS séma struktúrák, amelyek alkalmazhatóak kényszerek megadásában. Az osztály szerepe, hogy mechanizmusokat lehessen létrehozni RDF feldolgozó motorok számára, amellyel az RDF modell elemei értékelhetőek ki.

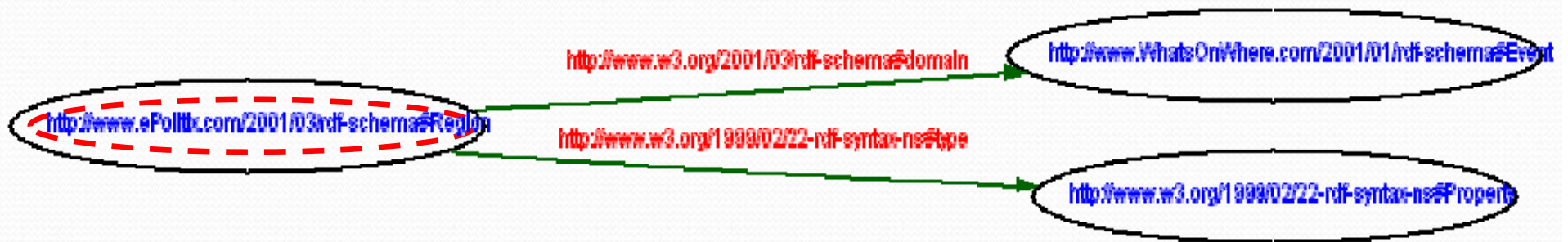
Tulajdonságok hasznosítása: hamis iskolai választás példa (RDF-statements)

```
<rdf:Description rdf:about="http://www.GrangeHill.edu/mark-birbeck">
  <rdf:type
    resource="http://www.Schemas.org/2001/01/rdf-schema#Person"/>
  <gh:Name>Mark Birbeck</gh:Name>
  <epx:CandidateFor
    rdf:resource="http://www.pa.press.net/constituencies/281" />
  <gh:Born>28 September, 1964</gh:Born>
</rdf:Description>
```



Tulajdonságok hasznosítása : Esemény web lap (egy RDFS részlet)

```
<rdf:Property rdf:about="http://www.ePolitix.com/2001/03/rdf-schema#Region">  
  <rdfs:domain  
    rdf:resource="http://www.WhatsOnWhere.com/2001/01/rdf-schema#Event" />  
</rdf:Property>
```



Több típus megadása

```
<rdf:Description
  rdf:about="http://www.epolitix.com/austin-mitchell"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:s="http://www.Schemas.org/2001/01/rdf-schema#"
  xmlns:epx="http://www.ePolitix.com/2001/03/rdf-schema#" >

  <rdf:type
    resource="http://www.Schemas.org/2001/01/rdf-schema#Person" />
    <s:Name>Austin Mitchell</s:Name>
    <s:DOB>19 September, 1934</s:DOB>
  <rdf:type
    resource="http://www.ePolitix.com/2001/03/rdf-schema#Candidate" />
    <epx:CandidateFor
      rdf:resource="http://www.pa.press.net/constituencies/281" />
    <epx:MemberOf>Labour Party</epx:MemberOf>
</rdf:Description>
```

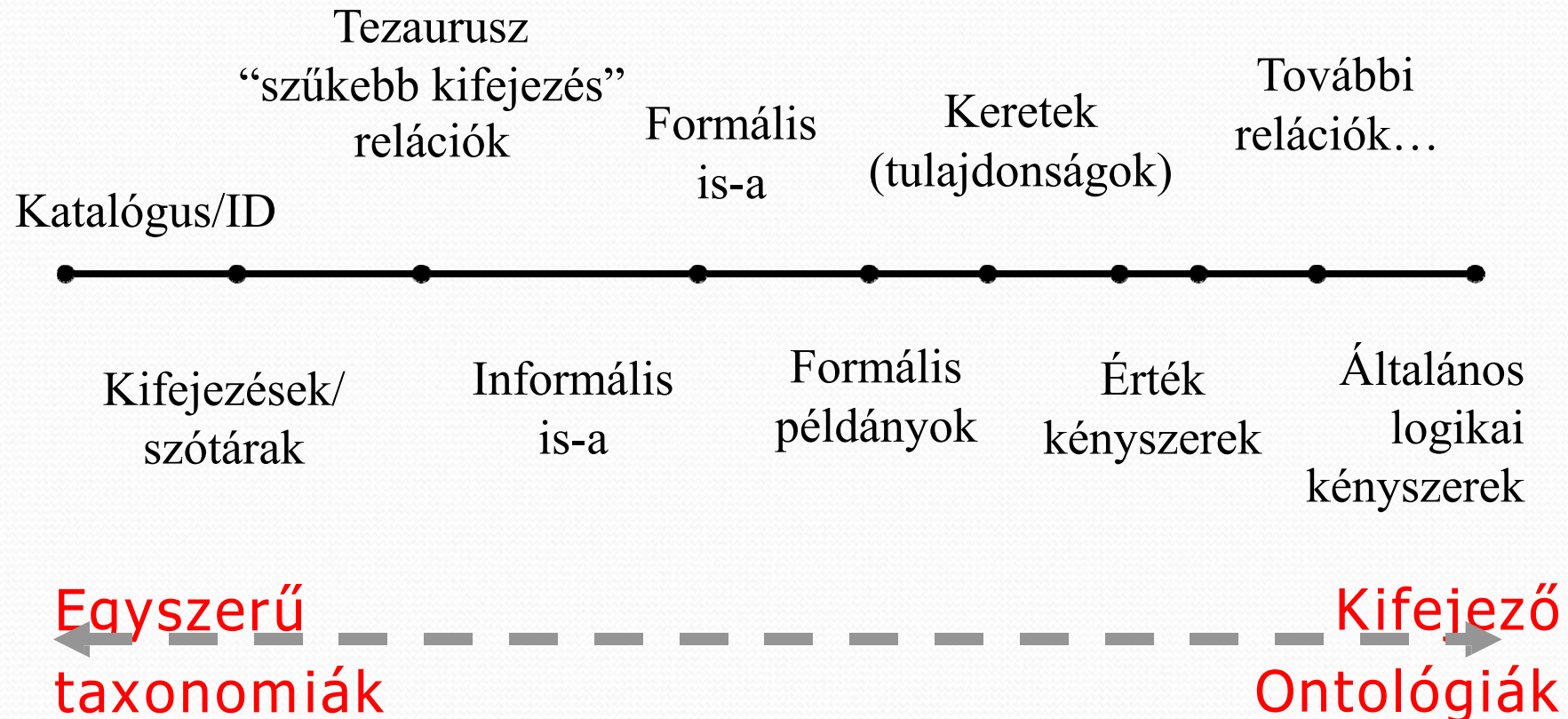
Több típus megadása



Osztályok az RDFS-ben

- Nincsenek metódusok, csak tulajdonságok
- Tulajdonság centrikus. OO világban az osztály metódusok és tulajdonságok halmazával definiált. RDF tulajdonságot az alkalmazhatóság (osztály) definiálja.
- Tulajdonságok osztályok kapcsolata nélkül is definiálható.

Szótárak, ontológiák



Ontológiák

- RDFS hasznos, de nem ad megoldást az elvárásokra, a szemantika pontos leírására.
- Összetett alkalmazások további igényei:
 - Tulajdonságok leírása, jellemzése
 - Különböző URI-val rendelkező objektumok azonosságának leírása (ekvivalencia)
 - Osztályok diszjunkt vagy éppen ekvivalens jellege
 - Osztályok konstruálása (nemcsak megnevezése)
 - Következtetési igények támogatása:
 - Pl.:“Ha két «Person» erőforrás «A» és «B» azonos «**foaf:email**» tulajdonsággal rendelkeznek, akkor «A» és «B» identikus.

Ontológiák

- Az SZW világban az ontológiákat a következő értelemben használjuk:

Fogalmak és ezek relációinak definiálása egy adott tudásterület leírása céljából.

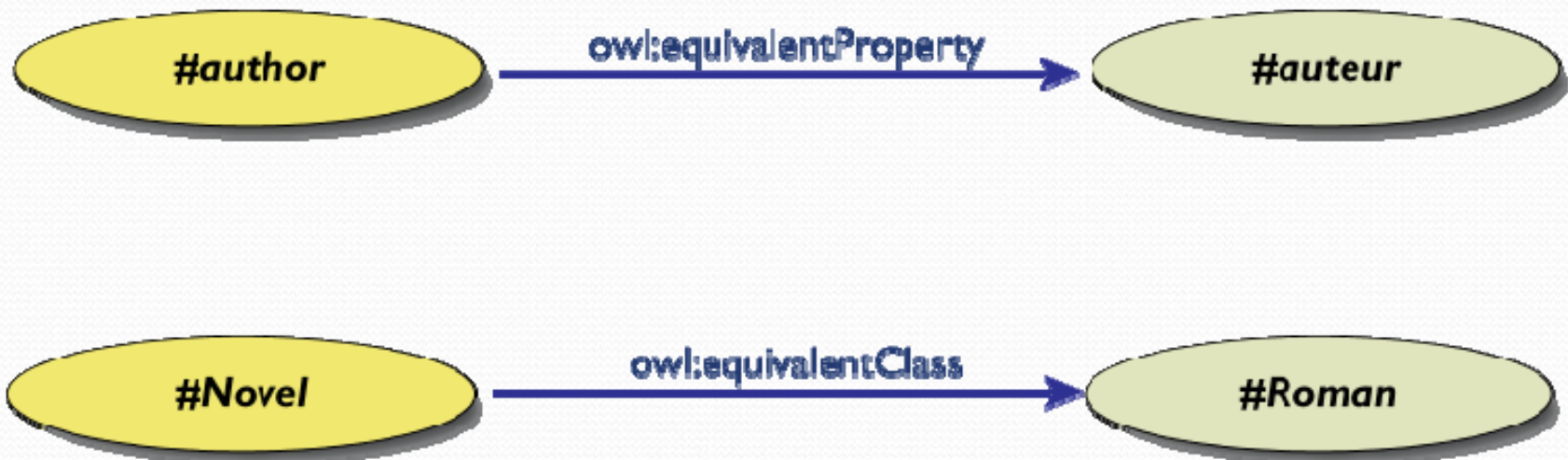
- Az RDFS is tekinthető egy egyszerű ontológia nyelvnek
- Nyelvek definiálása mindig egyfajta kompromisszum
 - gazdag szemantika tudás gazdag alkalmazásokhoz
 - ésszerűség (fizibilitás, megvalósíthatóság (fordító, következtető))

Web Ontology Language = OWL

- OWL az SZW struktúrába egy újabb réteg, az RDFS bővítése
 - Saját névterek, saját elemek, kifejezések
 - RDFS-re épül (tartalmazza)
- Önálló SZW ajánlás
 - “OWL 2” – 2010 óta gyakorlatilag csak ezt használjuk
 - Akit a részletek érdekelnek:
<https://www.w3.org/TR/owl2-overview/>

Példa: ekvivalencia relációk

- Osztályokra:
 - **owl:equivalentClass**: két osztálynak azonosak az elemei
 - **owl:disjointWith**: nincs közös elemük
- Tulajdonságokra:
 - **owl:equivalentProperty**
 - Példa: **a:author** vs. **f:auteur**
 - **owl:propertyDisjointWith**
- Egyedekre:
 - **owl:sameAs**: két URI ugyanazt a fogalmat vagy egyedet reprezentálja
 - **owl:differentFrom**: negált kifejezése az **owl:sameAs** kifejezésnek



Osztályok az OWL-ben

- RDFS: létező osztályokat alosztály struktúrába rendezhettük – osztályhierarchia építése, semmi több...
- OWL osztályok konstruálhatók más osztályok vagy példányok alapján, :
 - Elemek felsorolásával
 - Osztályok relációinak alkalmazásával: intersection, union, complement, stb.

Ontológiák építése

- Technológia adott, feladat ontológiák létrehozása
 - Szükséges a tárgyterület alapos ismerete, rendszerezési képesség
 - Vannak közösségek, ahol jelentős tapasztalatok vannak (pl. könyvtárosok, orvosi tudományok)
 - Az OWL csak egy eszköz az ontológiák formalizálására
 - Nagy méretű ontológiák építése gyakran történik elosztott, közösségi módszerekkel
- Ontológiák létrehozásának elsődleges célja a megosztás, az újrafelhasználhatóság biztosítása
 - névtér mechanizmusok által
 - explicit import után

Adatbázis analógia

- Ontológia axiómák analóg elemek az adatbázis sémákkal
 - Séma definiálja a struktúrát és a megkötéseket
- Ontológia példányok analóg elemek adatbázis tartalommal
- De vannak különbségek...

Adatbázis -vs- Ontológia

Adatbázis:

- Zárt világ feltételezés (CWA)
 - Hiányzó információ „hamis” értéket képvisel
- Egyedi megnevezés (unique name assumption, UNA)
 - Minden egyed neve különböző
- Séma kényszerekként értelmezhető

Ontológia:

- Nyílt világ feltételezés (OWA)
 - Hiányzó információ nem ismertnek feltételezett
- Nincs UNA
- Ontológiai axiómák implikációkként értelmezhetőek (következtetési szabályok alkalmazhatóak)
 - Maga után vonzás alkalmazása

Adatbázis -vs- Ontológia

- Ugyanaz a példa:
 - Individual: HarryPotter
 - Facts: hasFriend RonWeasley
 - hasFriend HermioneGranger
 - hasPet Hedwig
 - Individual: Draco Malfoy
- Lekérdezés: Draco Malfoy barátja-e HarryPotternek?
 - Adatbázis: Nem
 - Ontológia: Nem tudjuk
 - OWA (nem levezethető hogy nem létezik a kért reláció)

Adatbázis -vs- Ontológia

- Például, ha adottak tények/adatok:

Individual: HarryPotter

Facts: hasFriend RonWeasley

hasFriend HermioneGranger

hasPet Hedwig

Individual: Draco Malfoy

- Lekérdezés: Hány barátja van HarryPotternek?

- Adatbázis: 2

- Ontológia: Legalább 1

- Nincs UNA: Ron és Hermione lehet akár ugyanaz a személy

Adatbázis -vs- Ontológia

- Például, ha adottak tények/adatok:

Individual: HarryPotter

Facts: hasFriend RonWeasley
hasFriend HermioneGranger
hasPet Hedwig

Individual: Draco Malfoy

DifferentIndividuals: RonWeasley HermioneGranger

- Lekérdezés: Hány barátja van HarryPotternek?
 - Adatbázis: 2
 - Ontológia: Legalább 2
 - OWA: Nem tudjuk, hogy van-e még több itt meg nem nevezett barátja

Adatbázis -vs- Ontológia

- Például, ha adottak tények/adatok:

Individual: HarryPotter

Facts: hasFriend RonWeasley

hasFriend HermioneGranger

hasPet Hedwig

Types: hasFriend only RonWeasley or HermioneGranger

Individual: Draco Malfoy

DifferentIndividuals: RonWeasley HermioneGranger

- Lekérdezés: Hány barátja van HarryPotternek?
 - Adatbázis: 2
 - Ontológia: 2!

Adatbázis -vs- Ontológia

- Adjunk új tényeket a modellünkhöz:

Individual: Dumbledore

Individual: Fawkes

Types: Phoenix

Facts: isPetOf Dumbledore

- Adatbázis válasza?

- A beillesztést visszautasítja, a megadott kényszerek alapján:
A hasPet értékkészlete személy (individual); Dumbledore viszont egy varázsló (CWA)

- Ontológia következtetés válasza?

- Levezeti, hogy Dumbledore személy (értékkészlet korlátozás)
- Szintén levezeti, hogy Dumbledore egy varázsló (csak varázslónak lehet Phoenix típusú állata)

Adatbázis lekérdezés végrehajtása

- A séma nem határozza meg az eredményt
 - Az adatoknak a sémán értelmezett kényszereknek kell megfelelniük
- A lekérdezés megválaszolása modell ellenőrzési feladat
 - Adat megfelelések ellenőrzése
- Hatékonyan implementálható

Ontológia lekérdezés végrehajtása

- Ontológia axiómák meghatározó szerepet játszanak
 - A válasz tartalmazhat implicit levezethető tényeket
 - Megválaszolhatók koncepcionális és extenzionális kérdések is:
 - Például lehet-e egy muglinak egy Phoenix a háziállata?
- A lekérdezés megválaszolása tétel bizonyítási feladat
 - Pl. logikai vonzatok levezetése
- Akár NP teljes is lehet a kiértékelés
 - Praktikus implementációk ennél azért jobban teljesítenek

Mikor használjunk ontológiát?

- Ontológia alkalmazása hasznos lehet, ha
 - A séma nagy vagy bonyolult, illetve ha lekérdezési időben szükség van kiértékelésére
 - Következtetés használható a séma ellenőrzésére, kapcsolatok felderítésére
 - Nem lehetséges vagy nem praktikus teljes információs bázis létrehozása
 - Pl. összetett struktúrák vagy folyamatok leírása
 - Megengedhető a hatékonyság csökkentése
- Adatbázisok alkalmazása hasznos, ha
 - Séma mérete, összetettsége kezelhető, nem vizsgálandó lekérdezési időben (kompromisszum: paraméterezett megoldások)
 - Teljes információ elérhető
 - Pl. könyvelés, pénzügy, árukészletek
 - Teljesítmény garanciák szükségesek

Szemantikus web példa

- Legyen ismert ez az állítás:
 - Budapest Magyarország fővárosa
- Háttértudásunk segítségével levezethetjük:
 - Budapest Magyarország fővárosa ✓
 - Magyarország egy állam ✓
 - Budapest egy város ✓
 - Budapest Magyarországon van ✓
 - Debrecen nem Magyarország fővárosa ✗
 - Budapest nem Ausztria fővárosa ✗
 - Budapest nem egy állam ✗
 - ...

Szemantikus web példa

- RDFS szintig nem tudjuk leírni a következőket:
 - Minden államnak pontosan egy fővárosa van
 - Tulajdonság számosság
 - Minden város csak egy ország fővárosa lehet
 - Függvény tulajdonság
 - Egy város nem lehet egyben egy állam is.
 - Diszjunkt osztályok
 - ...

OWL alapú ontológia

- "Debrecen nem Magyarország fővárosa ❌
- Miért nem?
 - Országoknak pontosan egy fővárosa van,
 - Debrecen és Budapest nem ugyanaz a város

- OWL:

```
:capitalOf a owl:InverseFunctionalProperty .  
:Budapest :capitalOf :Magyarország .  
:Budapest owl:differentFrom :Debrecen .
```

```
Lekérdezés { :Debrecen :capitalOf :Magyarország . } → false
```


OWL alapú ontológia

- Budapest nem Ausztria fővárosa ✘
- Miért nem?
 - Egy város csak egy ország fővárosa lehet
 - Ausztria és Magyarország nem azonos

- Hasonlóan:

```
:capitalOf a owl:FunctionalProperty .  
:Budapest :capitalOf :Magyarország .  
:Ausztria owl:differentFrom :Magyarország .
```

```
Lekérdezés { :Budapest :capitalOf :Ausztria . } → false
```

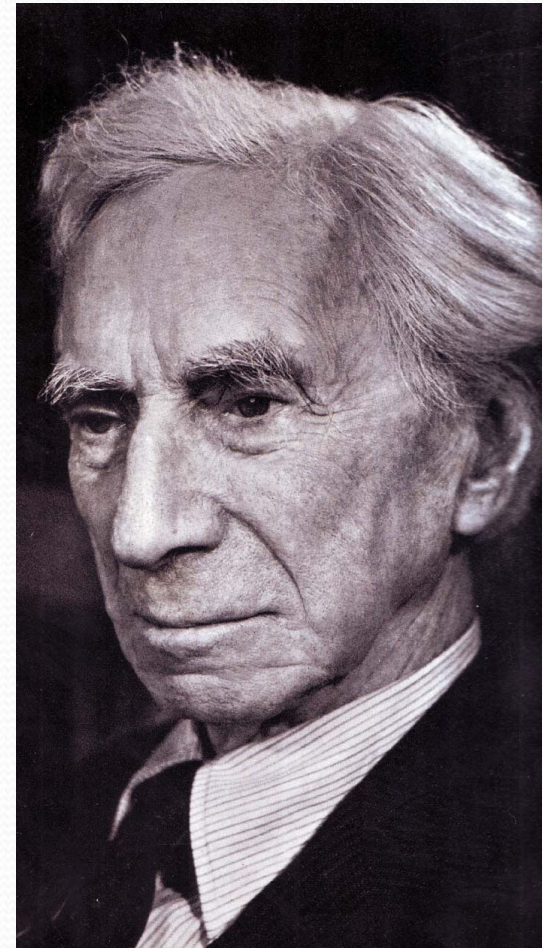
OWL2 példa: Russell paradoxon

Paradoxon

Bertrand Russell, 1918

Egy városban, ahol pontosan egy borbély van, mindenkit megborotvál, aki nem borotválja meg önmagát.

Ki borotválja meg a borbélyt?



OWL2 példa: Russell paradoxon

Osztály definíciók

```
:People owl:disjointUnionOf  
  (:PeopleWhoShaveThemselves  
  :PeopleWhoDoNotShaveThemselves ) .
```

Relációk:

```
:shavedBy rdfs:domain :People .  
:shavedBy rdfs:range :People .  
:shaves owl:inverseOf :shavedBy .
```

Szabály:

```
:People rdfs:subClassOf [  
  a owl:Restriction ;  
  owl:onProperty :shavedBy ;  
  owl:cardinality "1"^^xsd:integer ] .
```

OWL2 példa: Russell paradoxon

- Borbély definíciója:

- ```
:Barbers rdfs:subClassOf :People ;
owl:equivalentClass [
 • rdf:type owl:Class ;
 • owl:oneOf (:theBarber)
•] .
```



# OWL2 példa: Russell paradoxon

Magukat borotválók osztálya:

```
:PeopleWhoShaveThemselves owl:equivalentClass [
 rdf:type owl:Class ;
 owl:intersectionOf
 (:People
 [
 a owl:Restriction ;
 owl:onProperty :shavedBy ;
 owl:hasSelf "true"^^xsd:boolean
]
)
] .
```

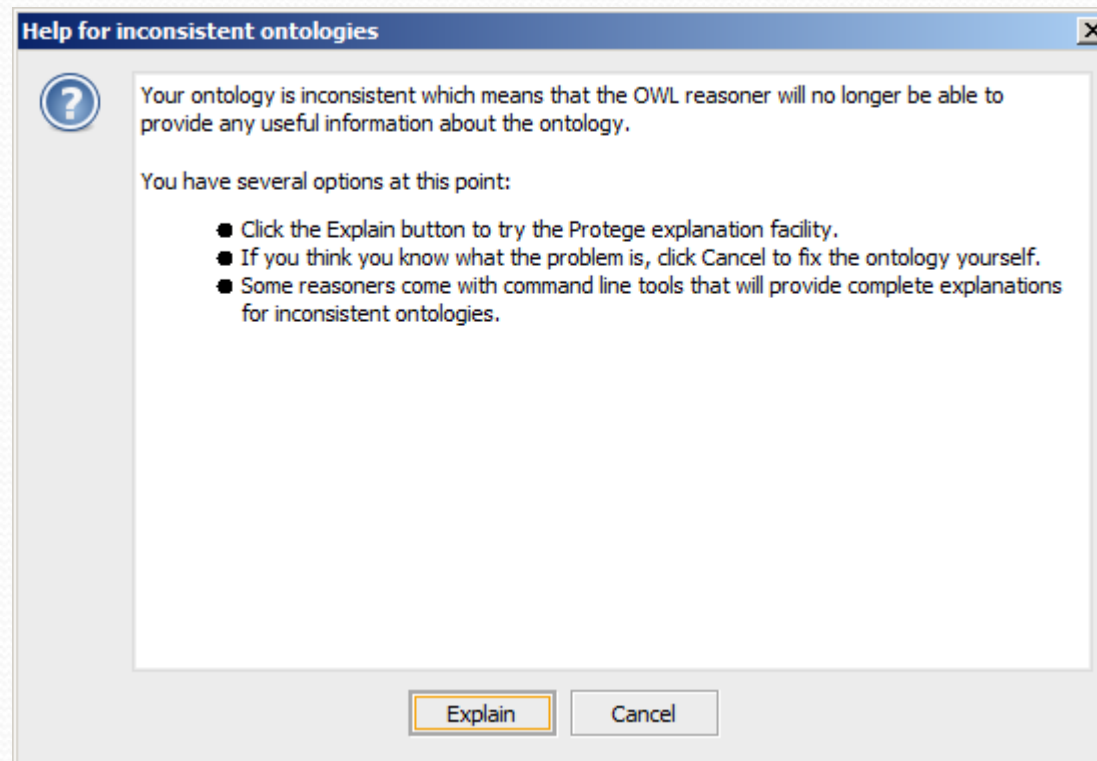
# OWL2 példa: Russell paradoxon

Akik nem magukat borotválják:

```
:PeopleWhoDoNotShaveThemselves owl:equivalentClass [
 a owl:Class ;
 owl:intersectionOf (
 :People
 [a owl:Restriction
 owl:onProperty :shavedBy ;
 owl:allValuesFrom :Barbers
]
)
] .
```



# OWL2 példa: Russell paradoxon



**Inconsistent ontology explanation** ✕

Show regular justifications     All justifications  
 Show laconic justifications     Limit justifications to

Explanation 1     Display laconic explanation

Explanation for: Thing SubClassOf Nothing

|    |                                                                                               |                             |   |
|----|-----------------------------------------------------------------------------------------------|-----------------------------|---|
| 1) | PersonsWhoDoNotShaveThemselves(?x) -> shaves(the-barber, ?x)                                  | In 1 other justifications   | ? |
| 2) | PersonsWhoDoNotShaveThemselves <b>DisjointWith</b> PersonsWhoShaveThemselves                  | In ALL other justifications | ? |
| 3) | Barber <b>SubClassOf</b> Person                                                               | In ALL other justifications | ? |
| 4) | shaves(?x, ?x) -> PersonsWhoShaveThemselves(?x)                                               | In ALL other justifications | ? |
| 5) | shaves(the-barber, ?x) -> PersonsWhoDoNotShaveThemselves(?x)                                  | In 1 other justifications   | ? |
| 6) | PersonsWhoShaveThemselves(?x) -> shaves(?x, ?x)                                               | In ALL other justifications | ? |
| 7) | Person <b>EquivalentTo</b> PersonsWhoDoNotShaveThemselves <b>or</b> PersonsWhoShaveThemselves | ALL other justifications    | ? |
| 8) | the-barber <b>Type</b> Barber                                                                 | In ALL other justifications | ? |

OK



# Szemantika

