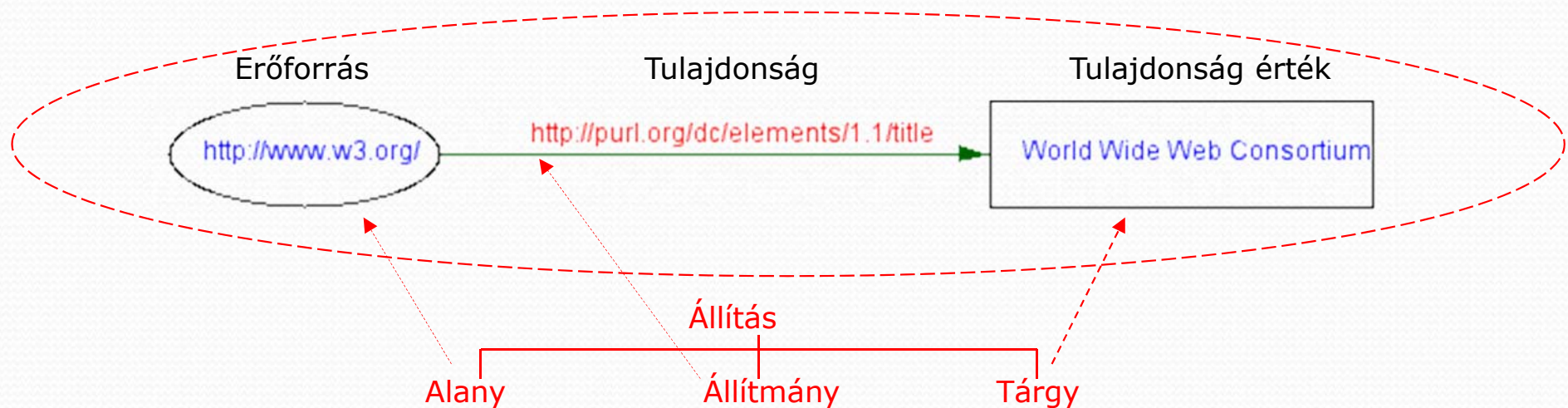


RDFS

(Resource Description Frameworks
Schema)

RDF

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

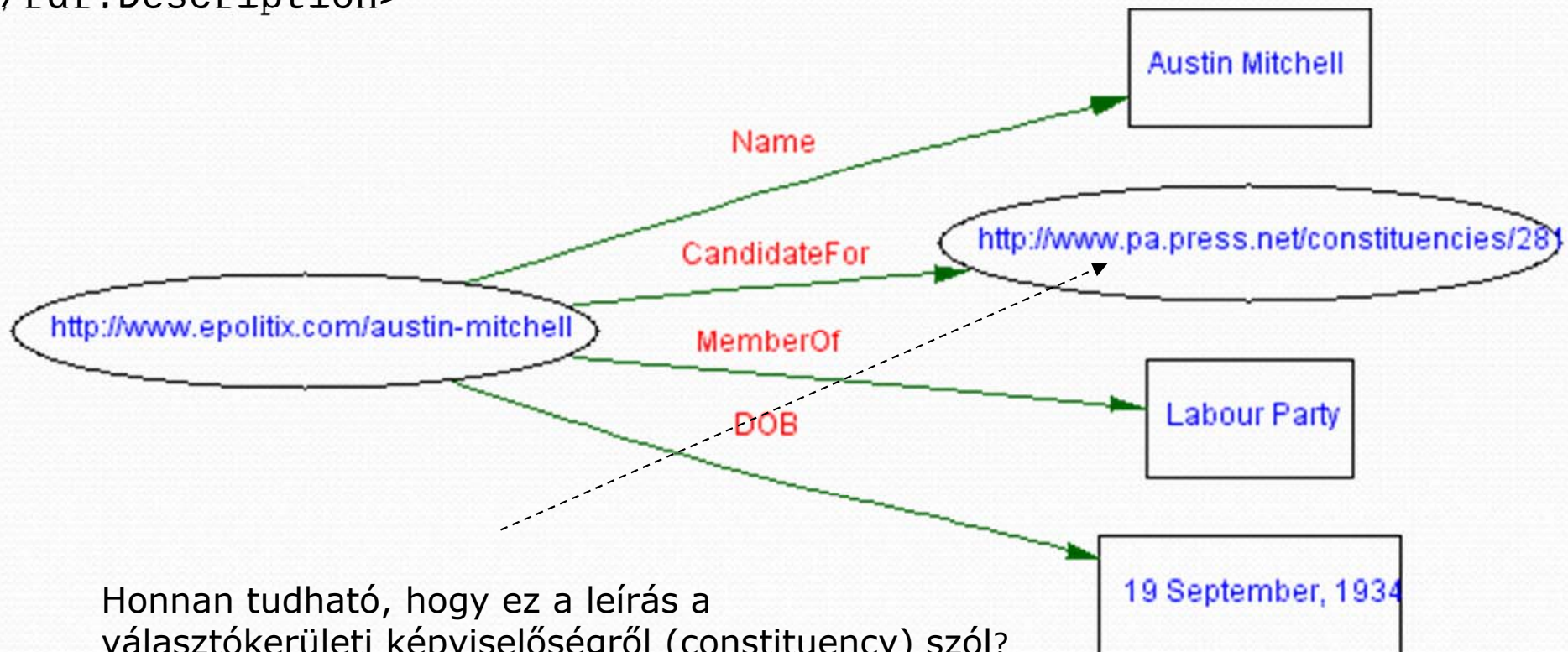


Mire használható az RDFS

- 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

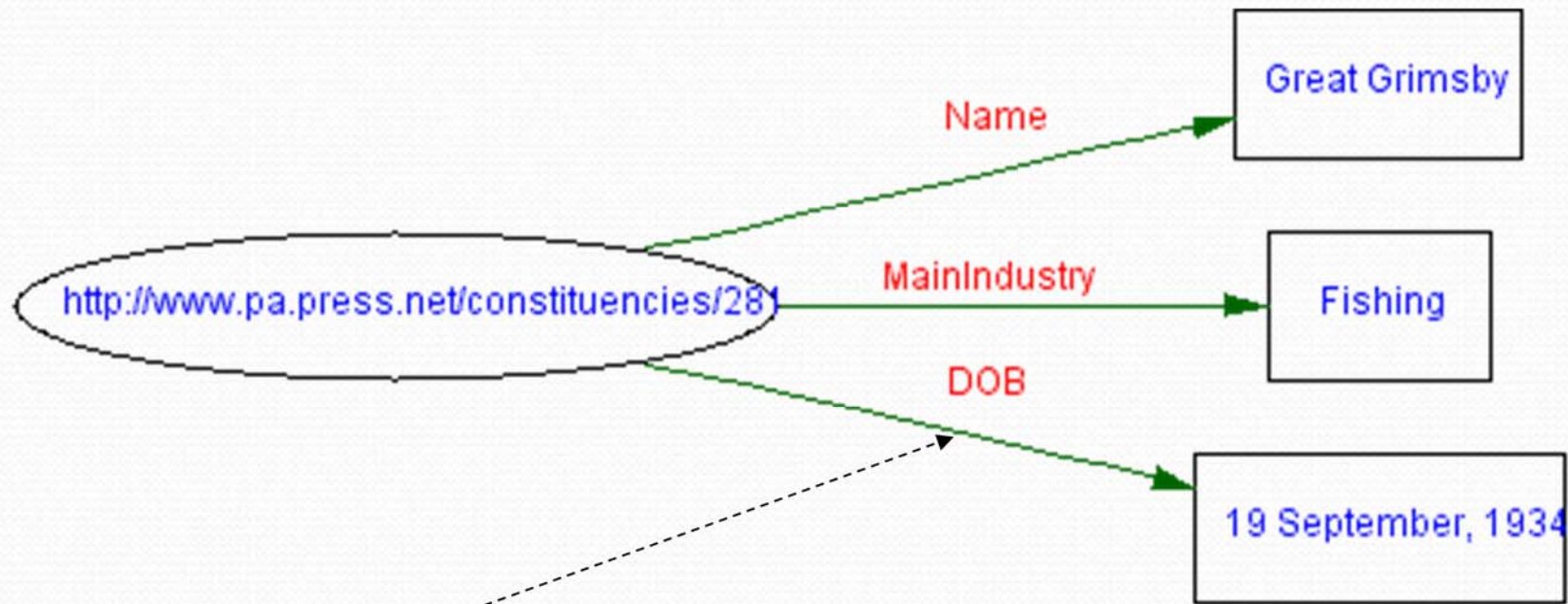
```
<rdf:Description rdf:about="http://www.epolitix.com/austin-mitchell">  
  <epx:Name>Austin Mitchell</epx:Name>  
  <epx:CandidateFor  
    rdf:resource="http://www.pa.press.net/constituencies/281" />  
  <epx:MemberOf>Labour Party</epx:MemberOf>  
  <epx:DOB>19 September, 1934</epx:DOB>  
</rdf:Description>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



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 választókörzet?

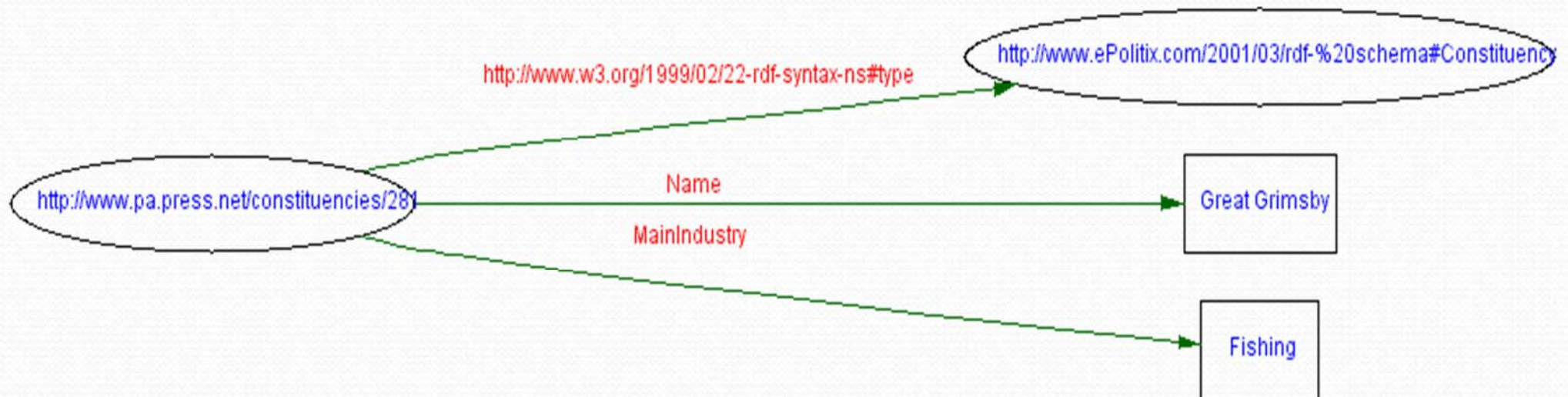
```
<rdf:Description rdf:about="http://www.pa.press.net/constituencies/281">  
  <epx:Name>Great Grimsby</epx:Name>  
  <epx:MainIndustry>Fishing</epx:MainIndustry>  
</rdf:Description>
```

- „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#" >

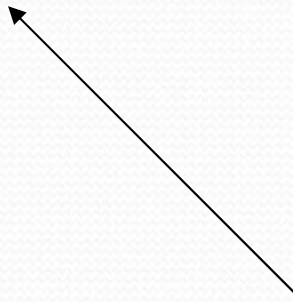
  <rdf:Description rdf:about=
    "http://www.pa.press.net/constituencies/281">
    <rdf:type resource=
      "http://www.ePolitix.com/2001/03/rdf-schema#Constituency"/>
    <epx:Name>Great Grimsby</epx:Name>
    <epx:MainIndustry>Fishing</epx:MainIndustry>
  </rdf:Description>
</rdf:RDF>
```



Tulajdonság értékek specifikálása: megfelelő állítmányok használata az alanyokhoz

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

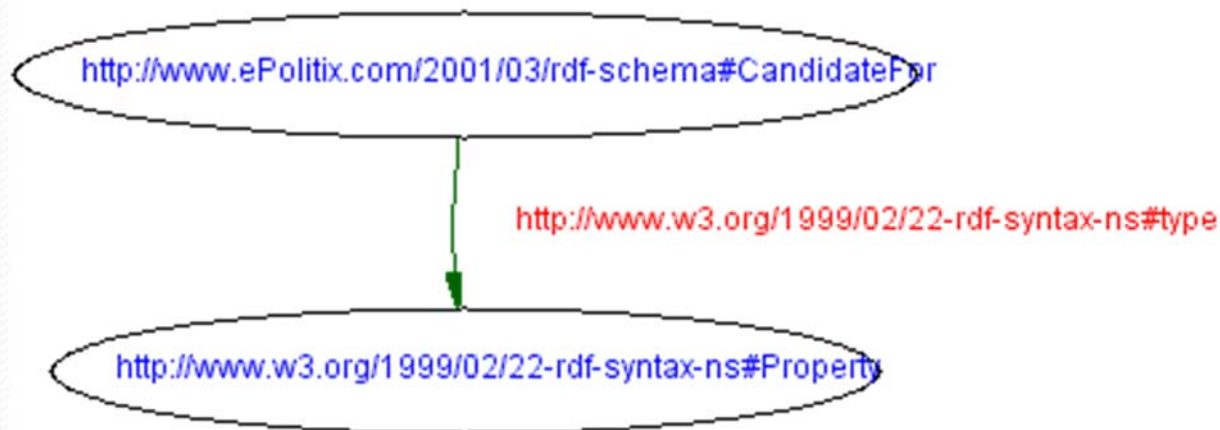
Érvénytelen: Nem választókerület



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.

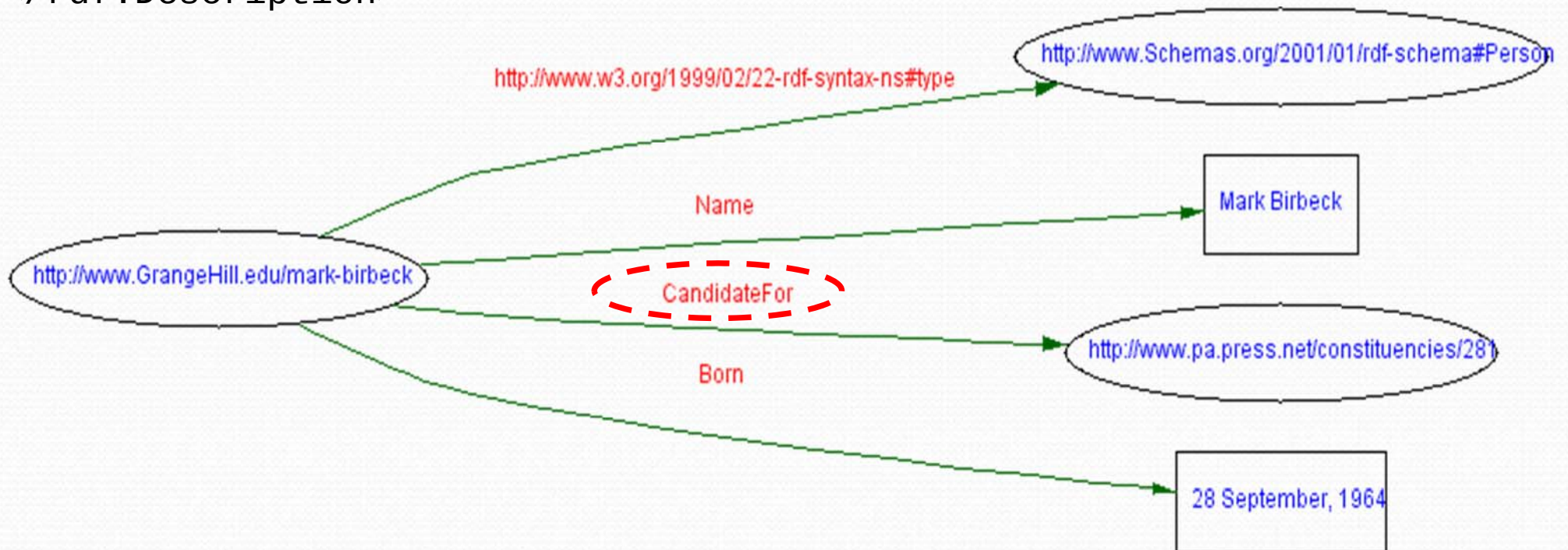
Az *rdfs:domain* és az *rdfs:range* példányai *rdfs:ConstraintProperty* osztálynak.

`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>
```

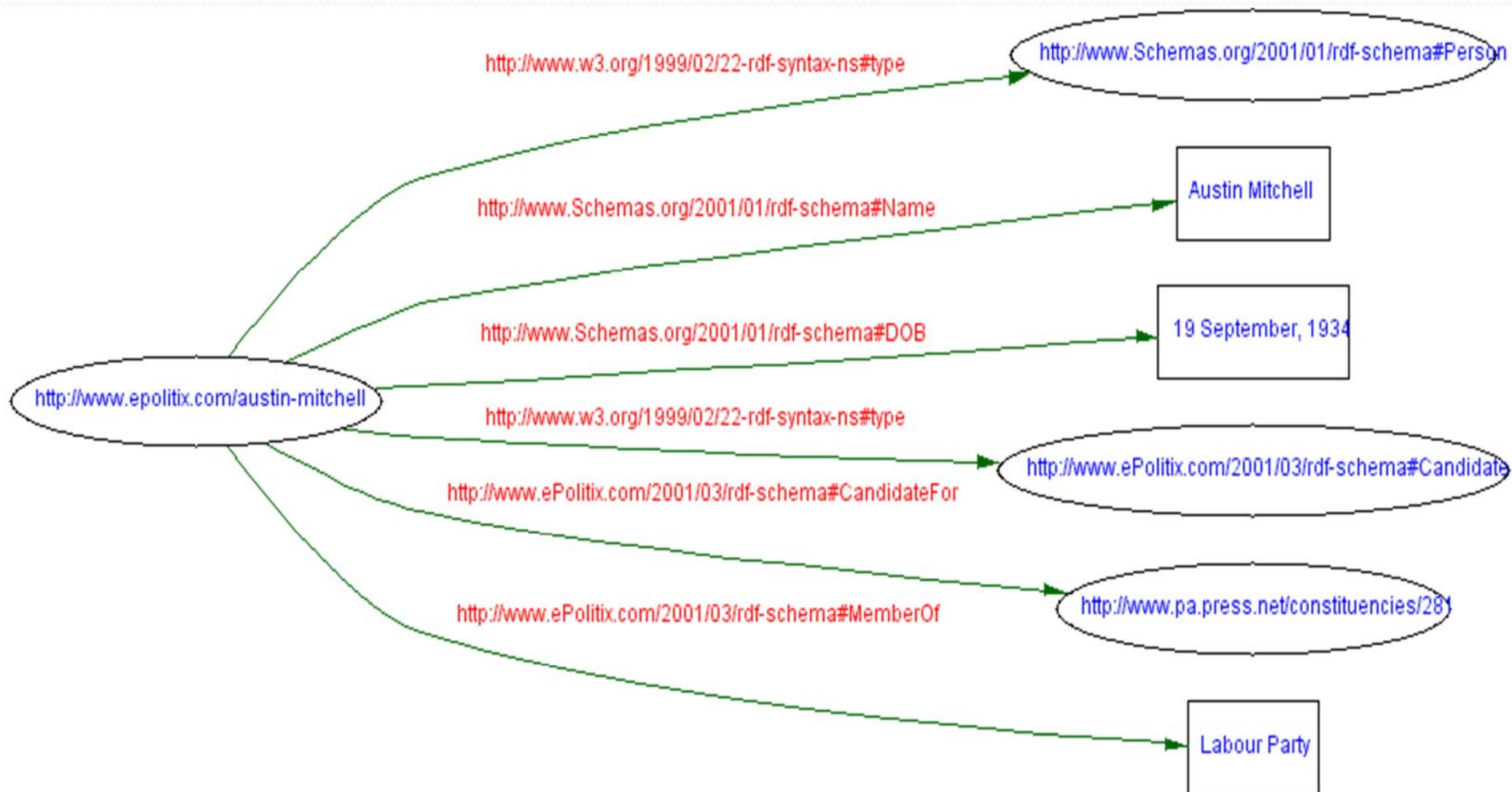


Több típus megadása

```
<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



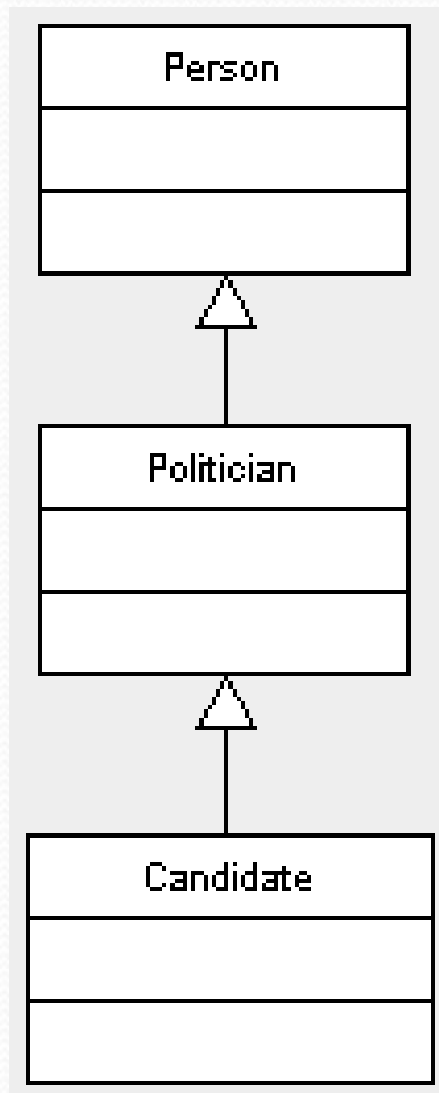
Osztályok az RDFS-ben

- 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.

Kategória rendszer építése: `rdfs:Class`

```
<rdfs:Class  
  rdf:about="http://www.Schemas.org/2001/01/rdf-schema#Person"  
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"  
>
```

Alosztályok létrehozása



Alosztályok létrehozása (XML szintaxis)

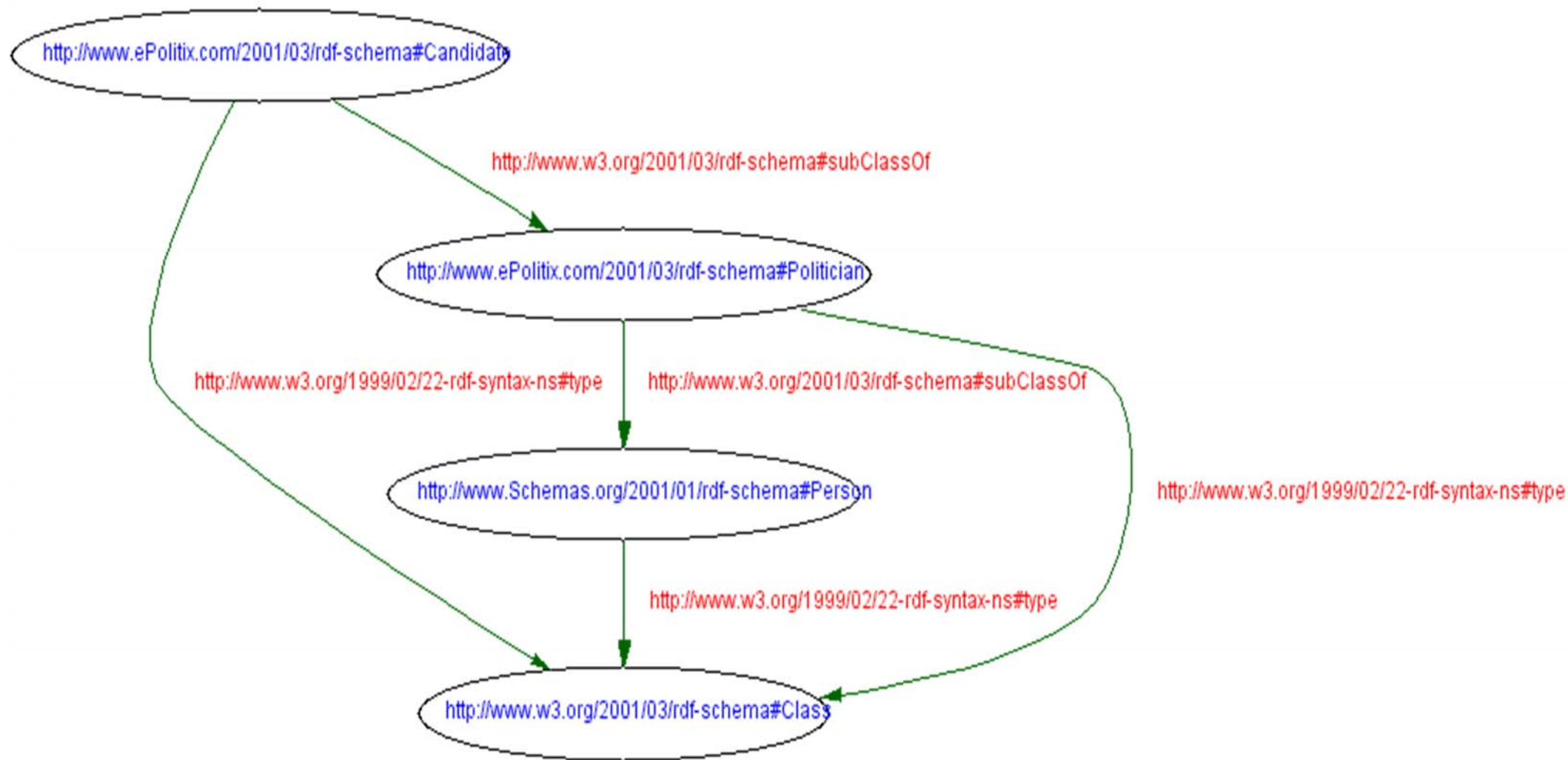
```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#" >

  <rdfs:Class rdf:about="http://www.Schemas.org/2001/01/rdf-schema#Person">

    <rdfs:Class
      rdf:about="http://www.ePolitix.com/2001/03/rdf-schema#Politician">
      <rdfs:subClassOf
        rdf:resource="http://www.Schemas.org/2001/01/rdf-schema#Person">
    </rdfs:Class>

    <rdfs:Class
      rdf:about="http://www.ePolitix.com/2001/03/rdf-schema#Candidate">
      <rdfs:subClassOf
        rdf:resource=
          "http://www.ePolitix.com/2001/03/rdf-schema#Politician" >
    </rdfs:Class>
  </rdf:RDF>
```

Alosztályok létrehozása

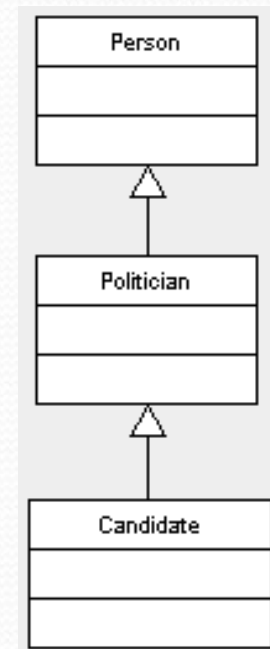


Alosztályok alkalmazása: RDFS Property

```
<rdf:RDF>
  <rdf:Property rdf:about="http://www.Schemas.org/2001/01/rdf-schema#Name">
    <rdfs:domain rdf:resource="http://www.Schemas.org/2001/01/rdf-schema#Person">
    <rdfs:range rdf:resource="http://www.w3.org/2000/01/rdf-schema#Literal">
  </rdf:Property>
  <rdf:Property rdf:about="http://www.Schemas.org/2001/01/rdf-schema#DOB">
    <rdfs:domain rdf:resource="http://www.Schemas.org/2001/01/rdf-schema#Person">
    <rdfs:range rdf:resource="http://www.w3.org/2000/01/rdf-schema#Literal">
  </rdf:Property>
</rdf:RDF>
```

```
<rdf:Property
  rdf:ID="MemberOf">
  <rdfs:domain rdf:resource="#Politician" >
  <rdfs:range rdf:resource="#PoliticalParty" >
</rdf:Property>
```

```
<rdf:Property
  rdf:ID="CandidateFor">
  <rdfs:domain rdf:resource="#Candidate" >
  <rdfs:range rdf:resource="#Constituency" >
</rdf:Property>
```



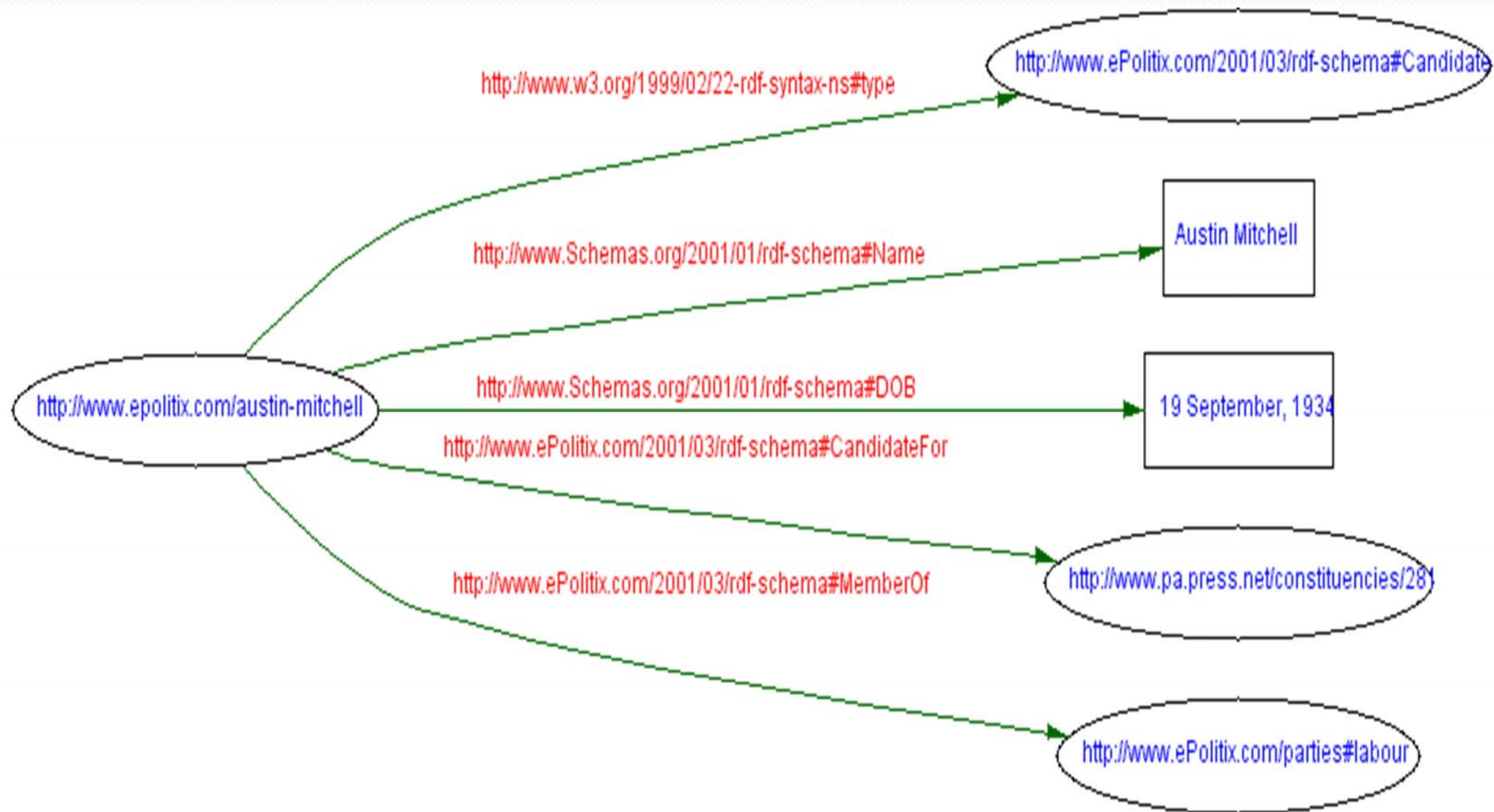
Alosztályok alkalmazása : RDF-statements

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

Alosztályok alkalmazása



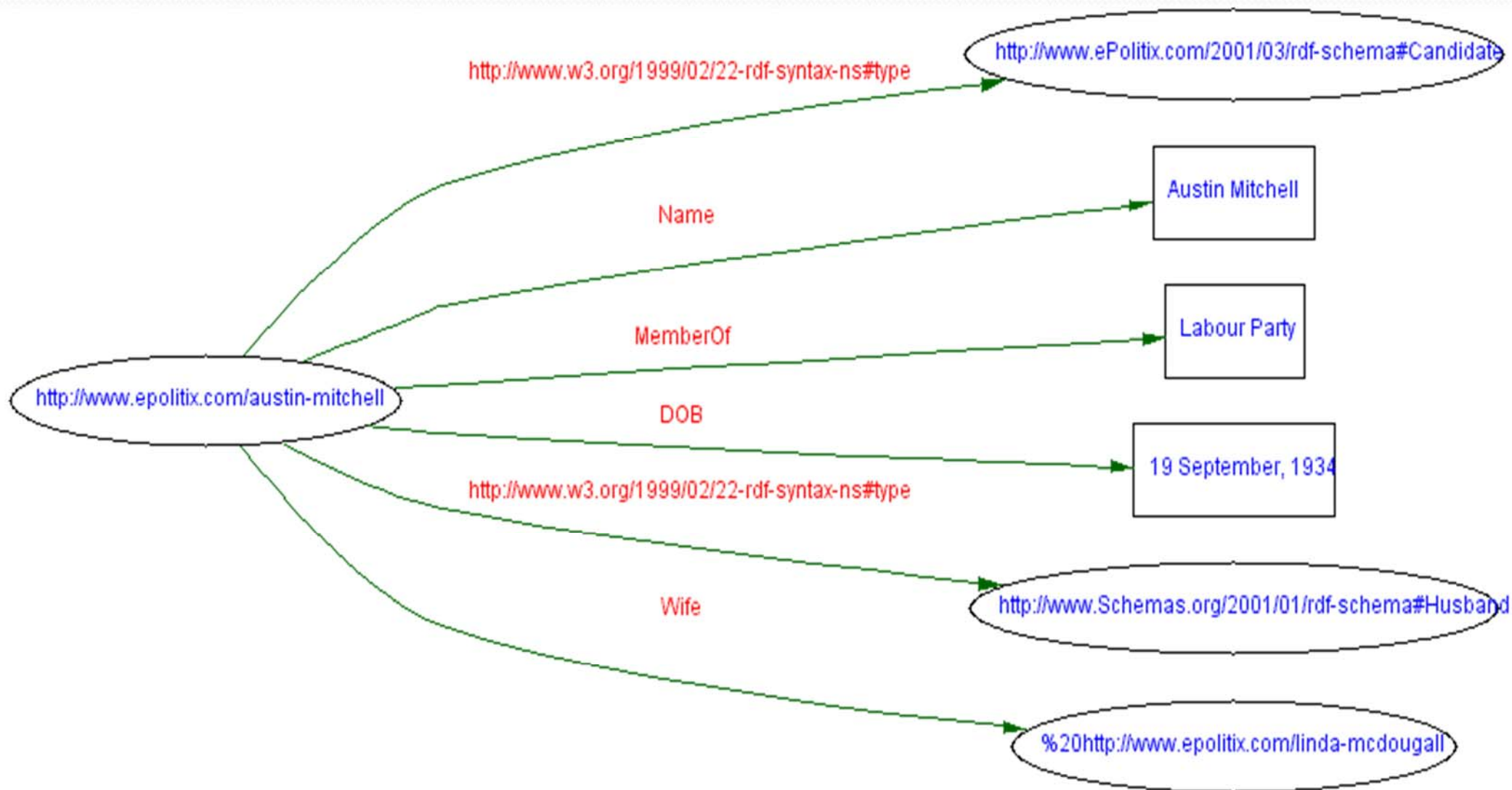
Többszörös öröklődés

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#" >

  <rdfs:Class
    rdf:about=
      "http://www.ePolitix.com/2001/03/rdf-schema#DefendingCandidate">
    <rdfs:subClassOf
      rdf:resource=
        "http://www.ePolitix.com/2001/03/rdf-schema#Candidate"/>
    <rdfs:subClassOf
      rdf:resource=
        "http://www.ePolitix.com/2001/03/rdf-schema#ElectedRepresentative" />
  </rdfs:Class>
</rdf:RDF>
```



Több példány



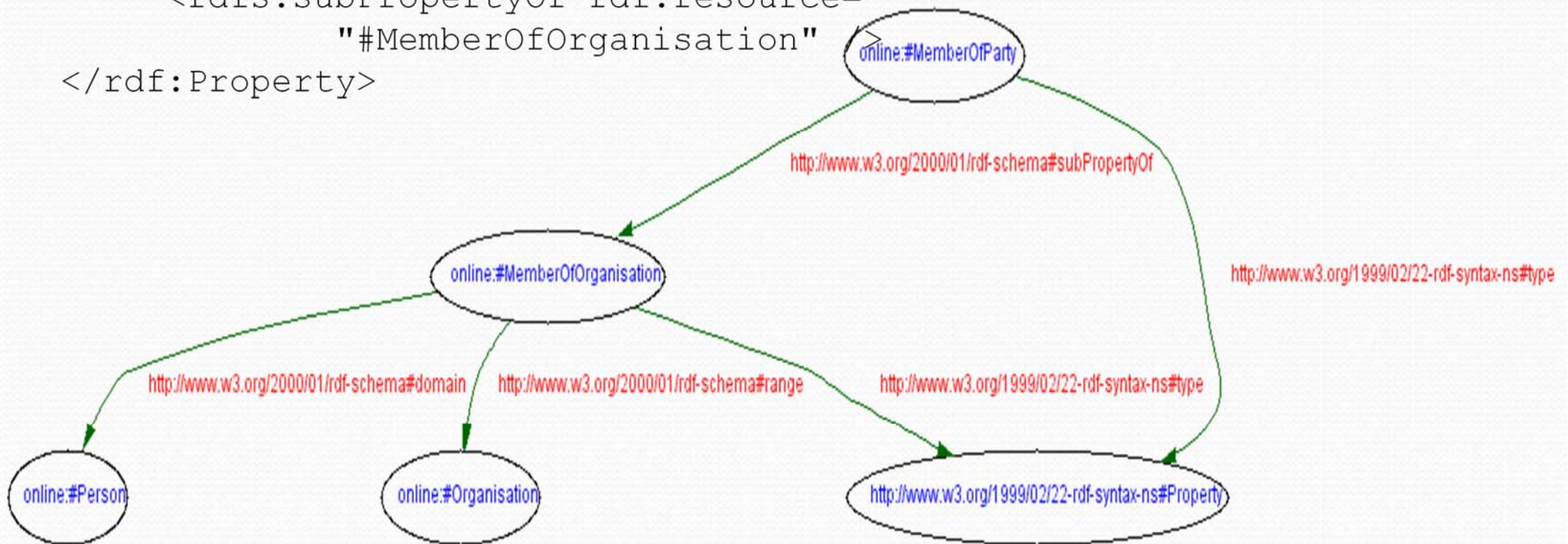
Több példány (XML szintaxis)

```
<rdf:Description
  rdf:about="http://www.epolitix.com/austin-mitchell">
  <rdf:type
    resource="http://www.ePolitix.com/2001/03/rdf-schema#Candidate"/>
  <s:Name>Austin Mitchell</s:Name>
  <epx:MemberOf>Labour Party</epx:MemberOf>
  <s:DOB>19 September, 1934</s:DOB>
  <rdf:type
    resource="http://www.Schemas.org/2001/01/rdf-schema#Husband" />
  <s:Wife rdf:resource=" http://www.epolitix.com/linda-mcdougall" />
</rdf:Description>
```

Résztulajdonságok (Subproperties)

```
<rdf:Property rdf:ID="MemberOfOrganisation">  
  <rdfs:domain rdf:resource="#Person" />  
  <rdfs:range rdf:resource="#Organisation" />  
</rdf:Property>
```

```
<rdf:Property rdf:ID="MemberOfParty">  
  <rdfs:subPropertyOf rdf:resource=  
    "#MemberOfOrganisation" />  
</rdf:Property>
```

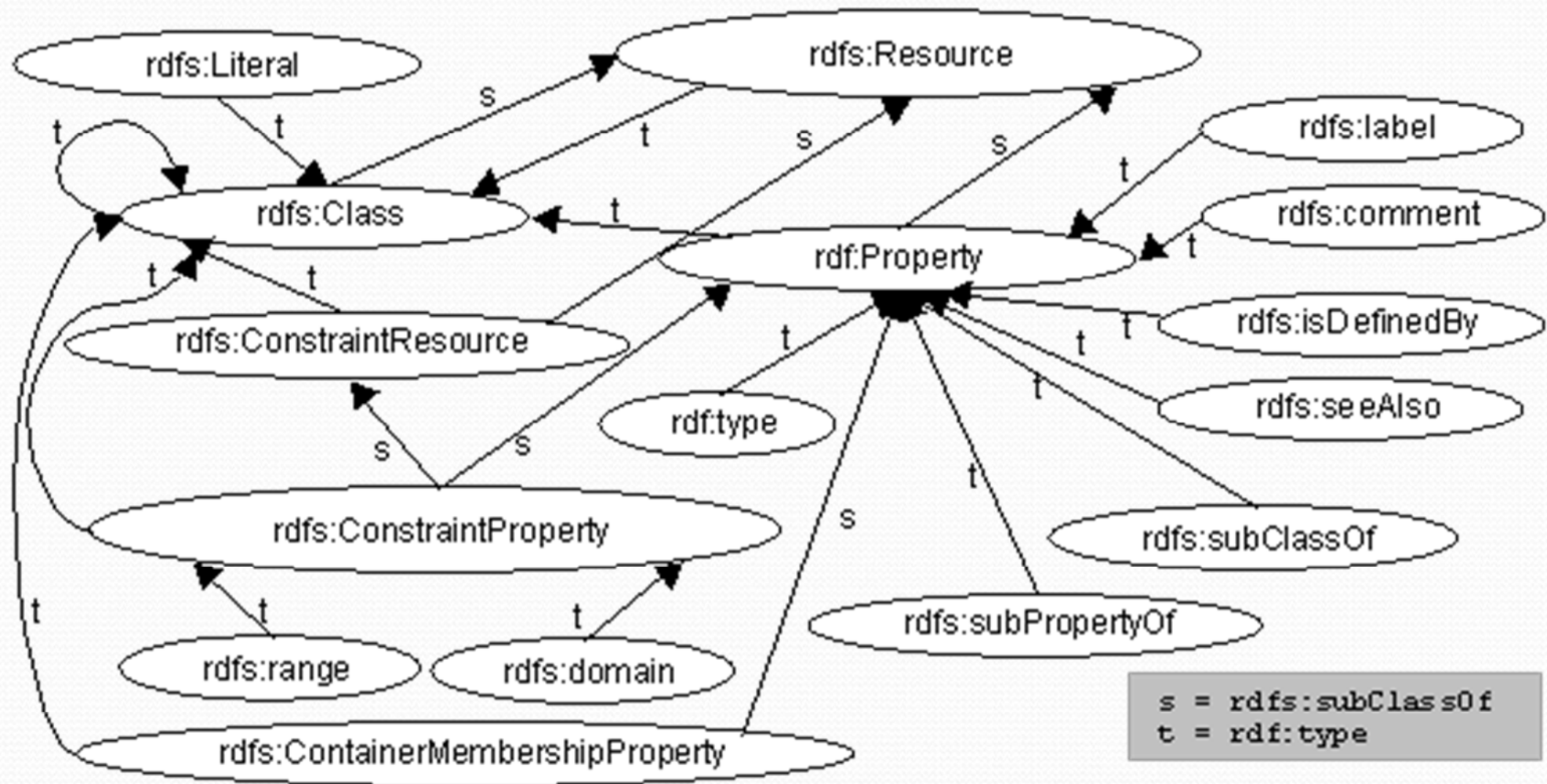


További RDFS elemek

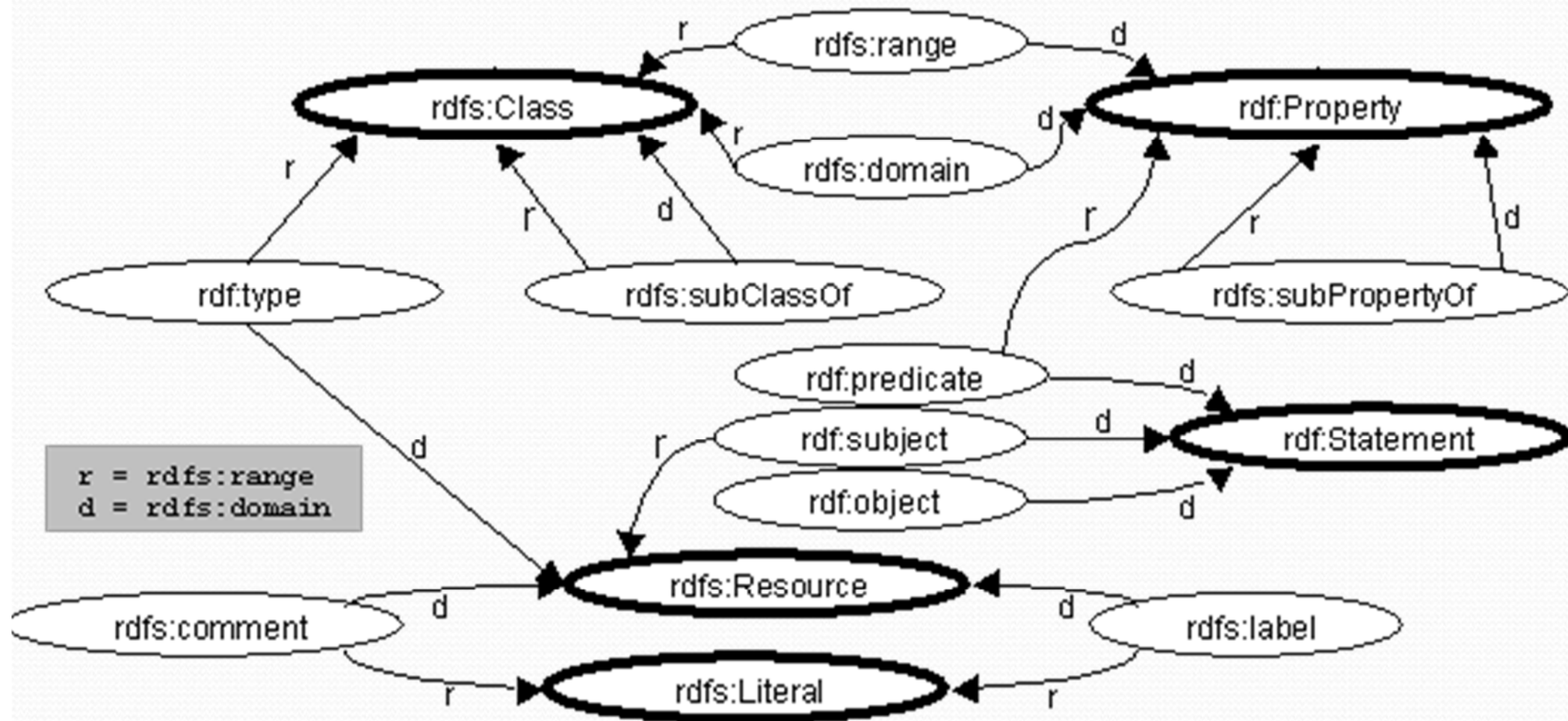
- `rdf:Resource`
- `rdfs:label`
- `rdfs:comment`
- `rdfs:seeAlso`
- `rdfs:isDefinedBy`
- `rdf:Statement`
- `rdf:subject`
- `rdf:predicate`
- `rdf:object`
- `rdfs:container`
- `rdf:bag`
- `rdf:seq`
- `rdf:alt`

Szabvány:
<http://www.w3.org/TR/rdf-schema/>

RDFS Osztály hierarchia



RDFS Kényszerek



Wordnet

(célok, motivációk)

Szótár alapú, pszicho-lingvisztikai alapok

- Egy jelentésalapú feldolgozása a lexikonoknak.
 - Fogalom alapú keresés egy lexikális adatbázison
- A fogalmakat szemantikus hálóba rendezzük
 - A lexikális információkat a szavak jelentése szerint rendezzük és nem a szavak formája szerint
- A Wordnet egy tezaurusz

A Wordnet szemantikus modellje

A „szó” egy asszociációs kapcsolat:

- a szótárba gyűjtött fogalmak, és
- a szó alakja között (szintaktika).

Lexikalizált fogalomgyűjtemény

Lexikális mátrix:

- Szó alakok(oszlopok)
- Szó jelentések (sorok).

Lexikális mátrix

Szinonimák

Szó jelentések	Szó alakok				
	F1	F2	F3 ...		F _n
M1	E1,1	E1,2			
M2		E2,2			
M3			E3,3		
...				...	
M _m					E _{m,n}

Többértelmű szavak

Reprezentációs módszer

- Konstruktív

- A reprezentáció tartalmazzon elegendő információt a fogalom felépítéséhez

- Differenciális

- A jelentések úgy legyenek fűzésekkel reprezentálva, hogy megkülönböztethetők legyenek

Wordnet fogalmak

Hipotézis:

- A *szinonima halmaz* egy megfelelő megközelítés egy fogalom definiálására.

Differenciális megközelítés

- A szavak jelentését reprezentálhatjuk egy szó-listával:
synset.

Angol wordnet tartalma

Tartalom: 95600 szóalak

- 51.500 egyszerű szó
- 44.100 kollokáció

70100 szó jelentés alak

Wordnet relációk

- Lexikális relációk (szóalakok között)
 - szinonimák
 - antonimák
- Szemantikus relációk (szó jelentések között)
 - Hiponima/Hiperonímia (alárendelt/fölérendelt)
 - Meronímia/Holonímia (része/magában foglalja)
 - Vonzat

Szinonima

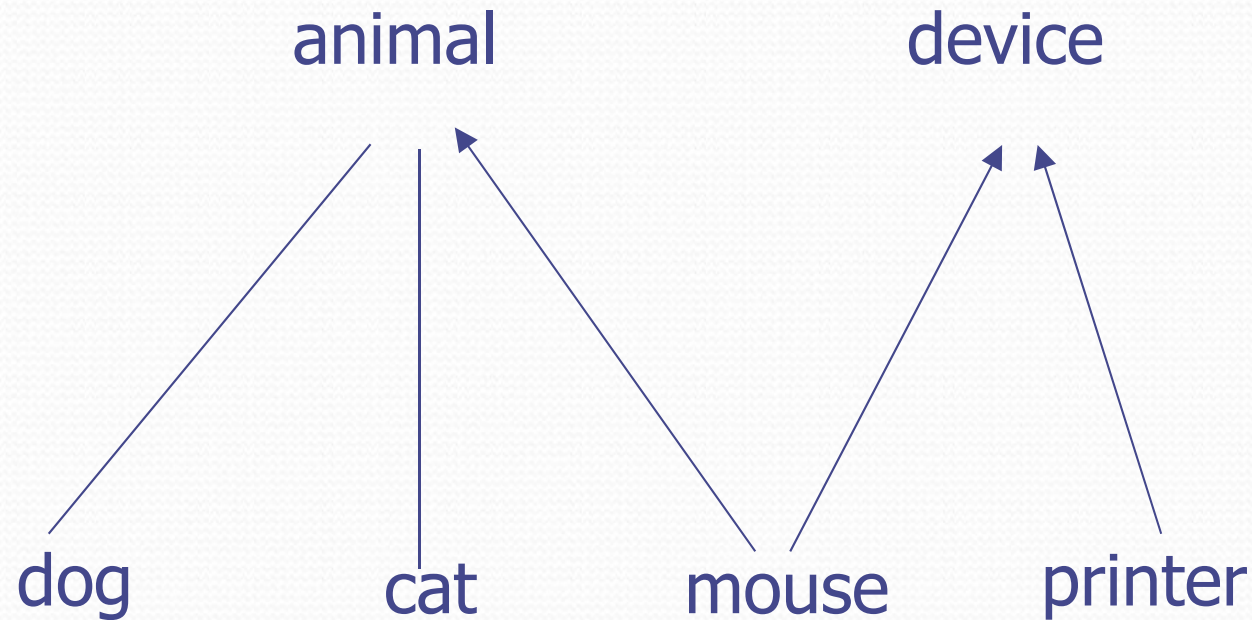
Két szó szinonima, azonos az értelmük, ha a következő kapcsolatok fennállnak:

- Minden szemantikus tulajdonságuk értéke megegyezik
- Ugyanannak a fogalomnak a megjelenései
- Kielégítik a Leibniz féle helyettesítési szabályt:
 - Ha felcseréljük a szinonimákat egy mondatban, akkor a mondat igazságtartalma nem változik

A Synset nem magyarázza el a foglom jelentését, de megjeleníti, igazolja a fogalom létezését

Hiponima

A hiponima egy olyan szókapcsolat, ahol a gyűjtő szó tartalmazza a kapcsolt szavak jelentését (alárendelés).



Meronima/Holonima

„Része” kapcsolat a jelentésben.

- Tranzitív és aszimmetrikus
- Egy tartalmazó fogalomhoz sok tartalmazott kapcsolódhat

Példa a Wordnet gazdagságára:

Rész-egész kapcsolat típusai a wordnet-ben:

- Component-object (branch/tree)
- Member-collection (tree/forest)
- Portion-mass (slice/cake)
- Stuff-object (aluminium/airplane)
- Feature-activity (paying/shopping)
- Place-area (Lausanne/Vaud)
- Phase-process (adolescence/growing up).

Szó kategóriák

Főnevek

- Hierarchiába szervezve – több (pl. hiponimák vagy meronimák szerint)

Igék

- Vonzat kapcsolatokon keresztül rendezve

Melléknevek

- Relációk (pl. ellentét) mentén rendezve

Kiindulópontok

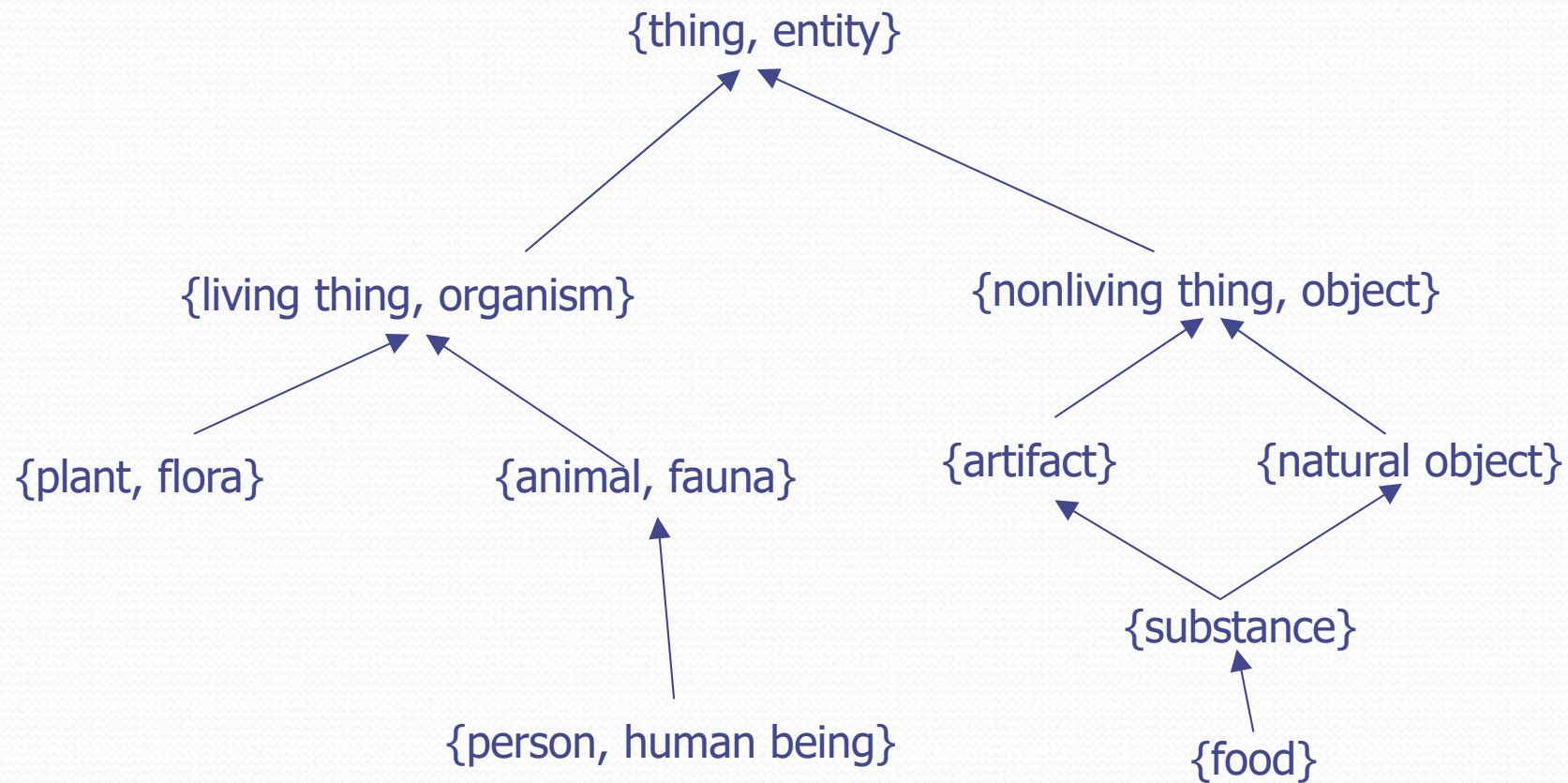
25 egyedi hierarchia

- Nem kölcsönösen kizáró kategóriák
- Keresztkapcsolatok megengedettek

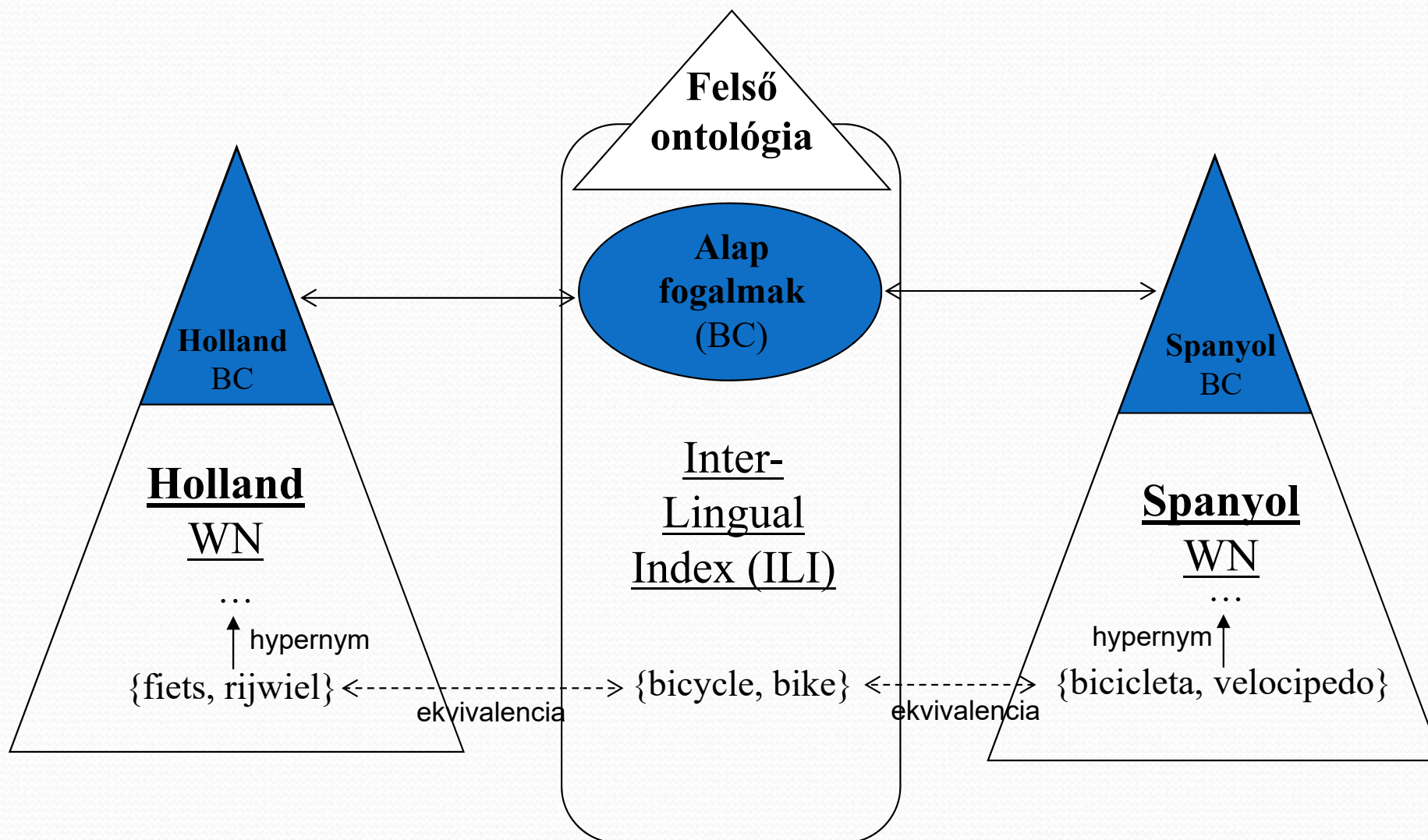
Fogalmi, lexikai területek:

{act, activity} {animal, fauna} {artifact} {attribute} {body} {cognition, knowledge} {communication} {event, happening} {feeling, emotion}	{food} {group, grouping} {location} {motivation, motive} {natural, object} {natural phenomenon} {person, human being} {plant, flora}	{possession} {process} {quantity, amount} {relation} {shape} {state} {substance} {time}
--	--	---

Alap kategóriarendszer a Wordnet-ben

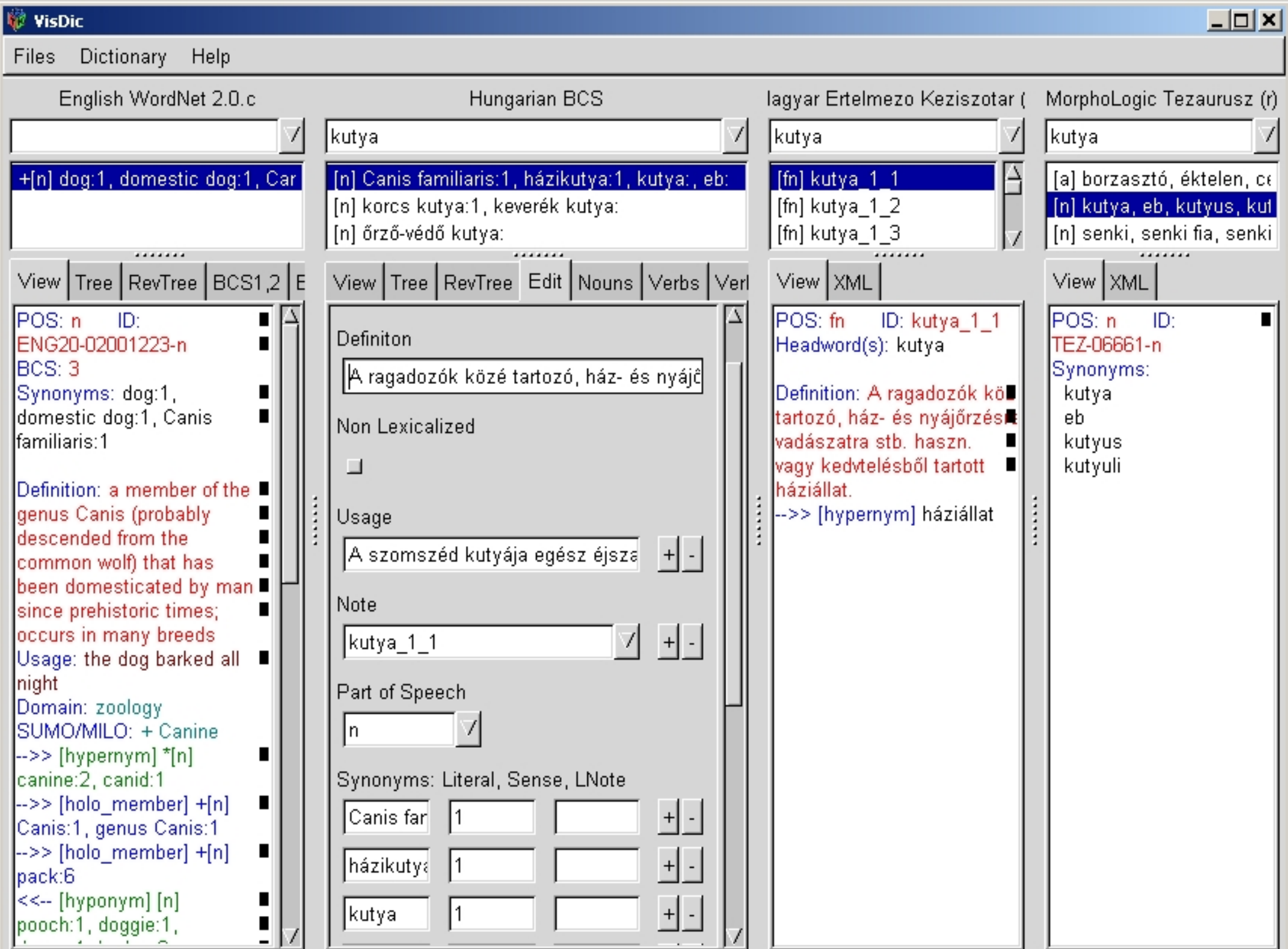


EuroWordNet: többnyelvű WN



Magyar WN ontológia (HuWN)

- **BalkaNet** projekt erőforrások használata
 - „Mag” rész: BN Concept Set (8 516 synset 13 nyelv alapján)
 - BN Interlingual Index (PWN 2.0 + SUMO hierarchia)
 - VisDic editor
- **Kiterjesztéses modell** (+ más)
 - Angol synsetek fordítása, relációk átvétele
 - Alapos kézi ellenőrzés és javítás
- **Fél-automatikus** módszerek
 - Korábban kifejlesztett fordító heurisztikák
 - 70% körüli pontosság (főnevek)
- Meglévő erőforrások **integrációja**
 - Magyar Értelmező Kéziszótár meghatározásai
 - NYTI igei vonzatkeret-adatbázis



HuWN: igék

- Problémák
 - Homályos jelentésbeli megkülönböztetések
 - Inkonzisztens angol WN
Thematikus szerepek, metaforikus jelentések, szelekciós megkötések stb.
- Megoldás
 - „Vegyes” metodológia:
BCS fordítás + MNSZ vonzatkeret-gyakoriság alapján kiválasztott igék, saját rendezés
 - Specifikus magyar relációk
Igekötők, -képzők kezelése stb.

HuWN

- Mag rész kiterjesztése
 - MNSZ és ÉKSz korpuszgyakoriságok alapján
- Ontológia további bővítése
 - Főnevek, melléknevek:
 - Iteratív koncentrikus bővítés PWN alapján
 - ÉKSz-ben feltárt szemantikai relációk alapján
 - Igék:
 - MNSZ vonzatkeret-gyakoriság alapján
 - PWN alapján
 - 2007: kb. 40K synset