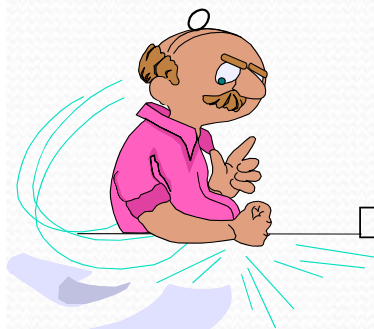


Intelligens elosztott rendszerek  
2022. tavasz

# SZEMANTIKUS WEB

*Méréstechnika és Információs Rendszerek Tanszék*

Probléma felvetés



Lekérdezés

Válasz

*Forrás leírás*

*Forrás fúzionálás/  
Lekérdezés tervezés*

*Végrehajtás*

**Mediátor**

*Monitor*

Web lapok

Strukturált  
adatok

Szenzorok  
(soros adatok)

# 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 „kereket”
- 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](mailto:ivan@w3.org)

# Böngészés a webben

Ivan Herman - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://www.w3.org/People/Ivan/

W3C

## Ivan Herman

[My Work at W3C](#) | [Contact information](#) | [Short CV](#) | [Upcoming trips](#) | [Public presentations](#)

### My Work at W3C

I am [Semantic Web Activity Lead](#); that is my main work at W3C. I am member of [IW3C2](#) (International World Wide Web Conference Committee) (the committee coordinating the yearly WWW conference series), serving as a liaison for W3C, and of [SWSA](#) (Semantic Web Science Association), the committee responsible for the International Semantic Web Conferences series.

As part of my work, I also participate in lots of outreach activities, and I regularly make presentations, tutorials, etc. You can consult my [list of presentations](#) for further details.

### Contact information

Email:  
[ivan@w3.org](mailto:ivan@w3.org)  
(sha1sum: 5ac8022d5f8042aa1775ca2f80e1676bafd5e80b)


Postal address:  
[C/o Centre for Mathematics and Computer Sciences \(CWI\)](#)  
[Kruislaan 413, P.O. Box 94079, 1090 GB Amsterdam, The Netherlands.](#)

Phone numbers:  
phone: +31-20-5924163  
mobile phone: +31-641044153  
fax: +31-20-5924312

PGP/GPG:  
My GnuPGP key and signature is available on-line.

FOAF:  
You can either extract a short FOAF information from this page, or consult my more complete, public FOAF file.

Misc:



# ... átugorva a munkahely linkre:

The screenshot shows a Mozilla Firefox browser window displaying the homepage of Centrum Wiskunde & Informatica (CWI). The browser's address bar shows the URL <http://www.cwi.nl/>. The website header features the CWI logo and navigation links for home, contact, a/z, intranet, and nl, along with a search bar. A main navigation menu includes links for about cwi, events, library, news, research, and scientists. The main content area is divided into several sections:

- Centrum Wiskunde & Informatica**: A brief description of the institute's research focus in mathematics and computer science, mentioning its strategy for 2012.
- Research Themes**: Four thematic areas are highlighted with images: Earth & life sciences, The data explosion, Societal logistics, and Software as service.
- News**: Two news items are listed:
  - 13-10-08 PhD defence Jarek Byrka**: Accompanied by a small image of a person.
  - 22-09-08 Best Paper Award for SMIL State research**: A news item about researchers Jack Jansen and Dick Bultman receiving an award at the ACM DocEngineering Symposium in Sao Paulo, Brazil.
- Agenda**: A list of upcoming events:
  - 18-10-08 Science Day at the Science Park Amsterdam**: A detailed announcement about the Science Day event, including the theme 'Crack the code' and a link to [read more](#).
  - 21-10-08 MAS Seminar, speaker Svetlana Dubinkina**: A seminar announcement with details on the speaker, session, and room.
  - 28-10-08 MAS Seminar, speaker Peter Sonneveld**: A partial announcement for another seminar.

# 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*





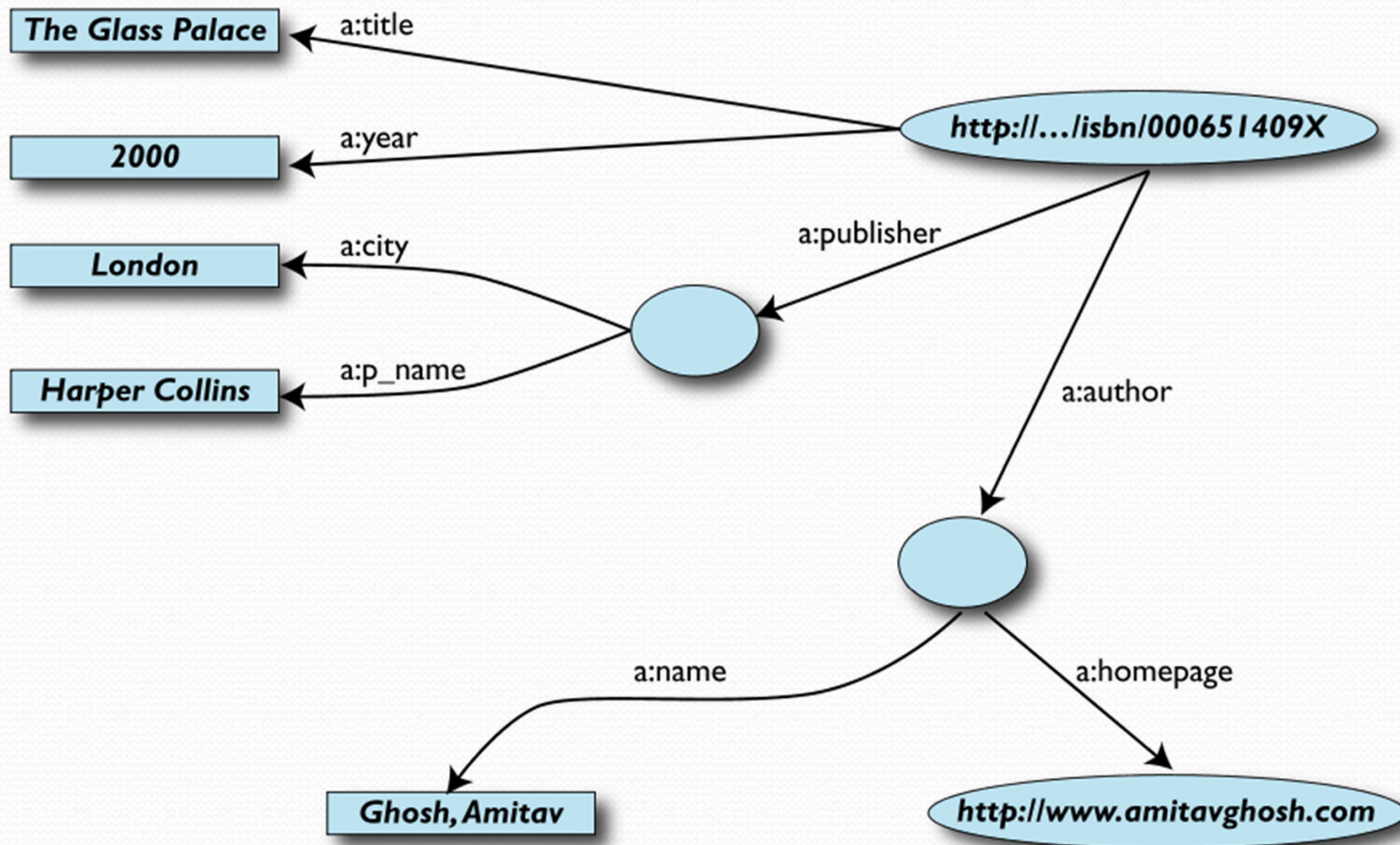
# Példa: könyvesbolt adatai

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

<b>ID</b>	<b>Name</b>	<b>Home Page</b>
id_xyz	Ghosh, Amitav	<a href="http://www.amitavghosh.com">http://www.amitavghosh.com</a>

<b>ID</b>	<b>Publ. Name</b>	<b>City</b>
id_qpr	Harper Collins	London

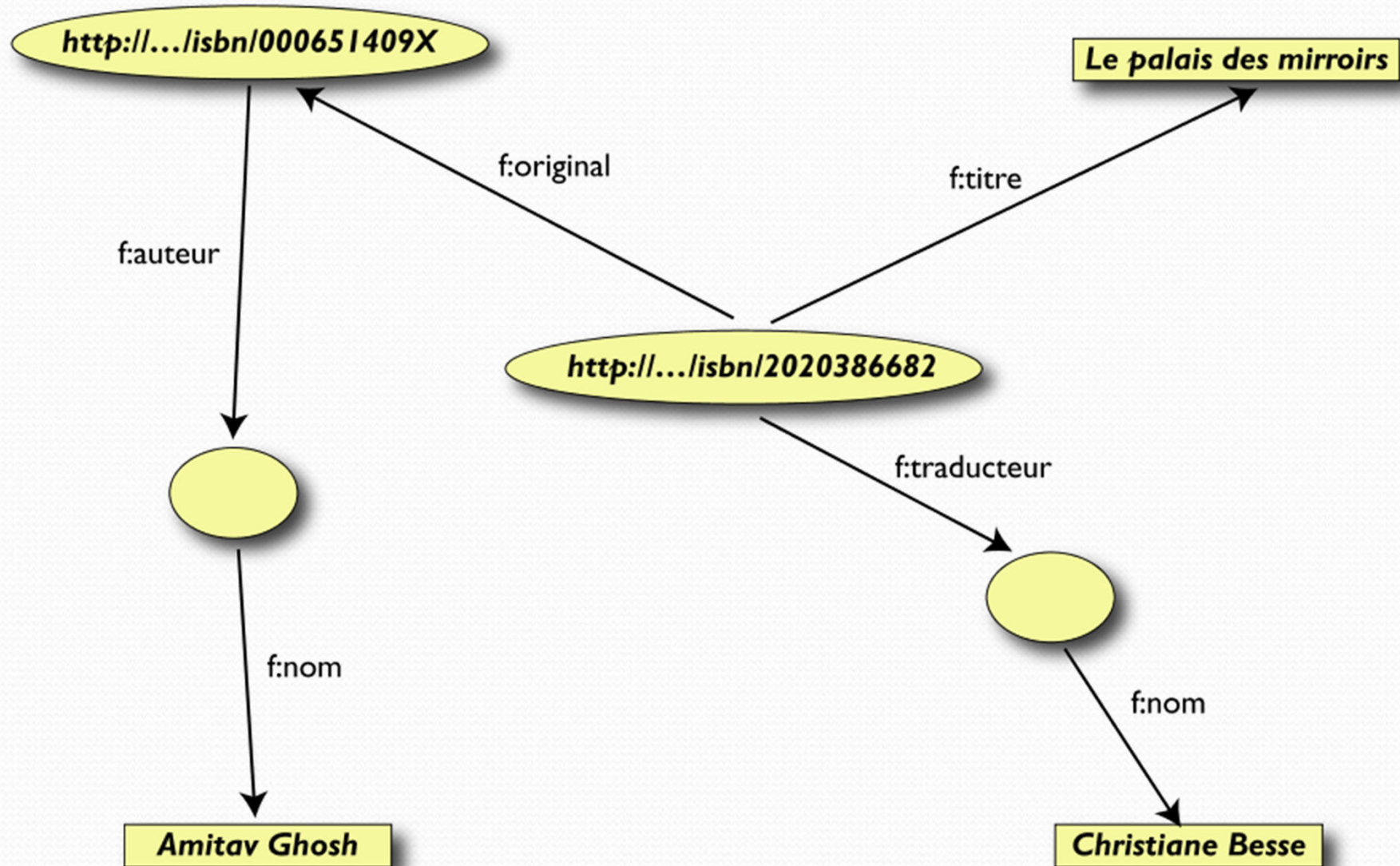
# Adatok exportálása relációkként



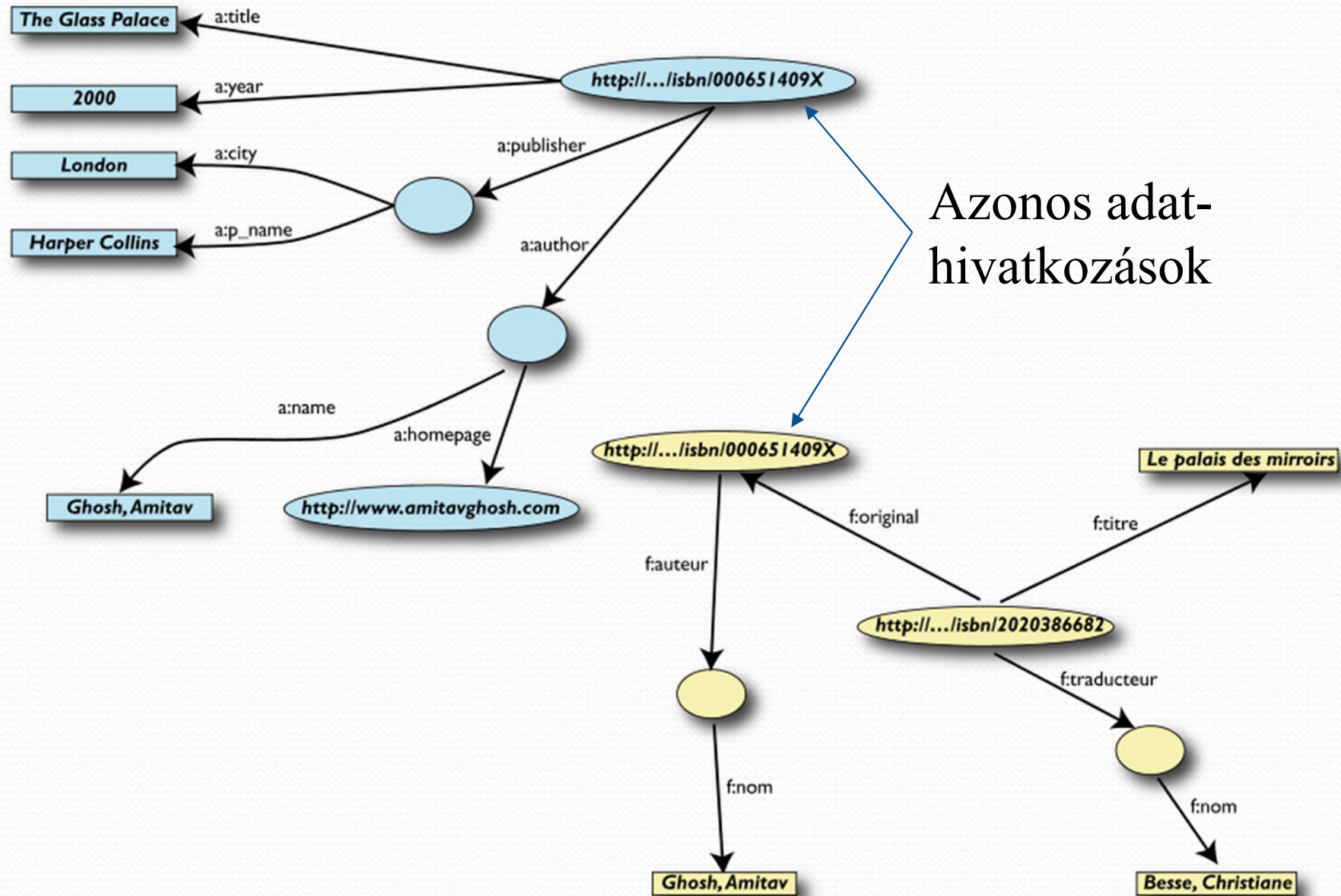
# Egy másik könyvesbolt adatai

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

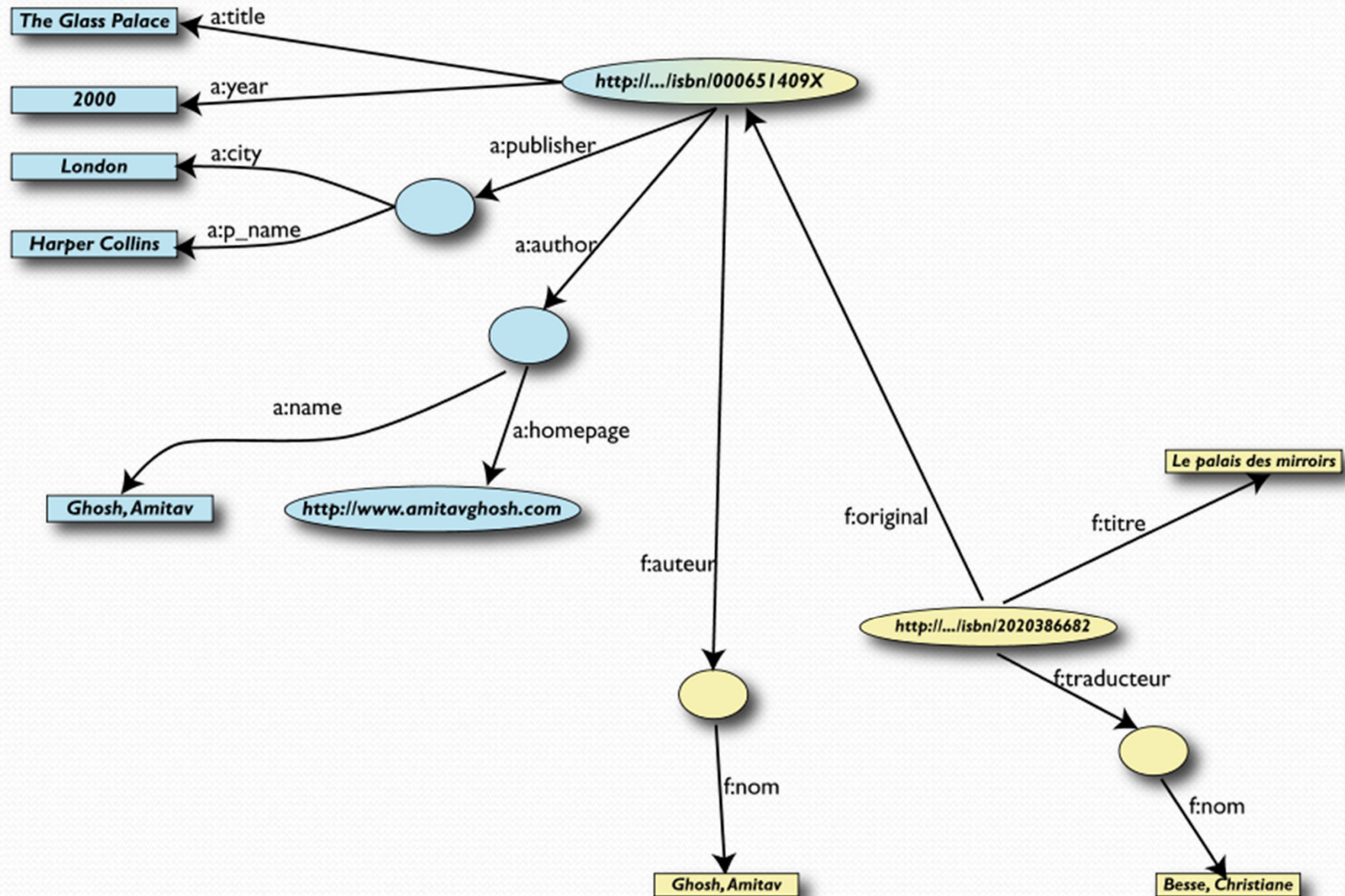
# A második könyvesbolt adatainak exportja



# Kapcsoljuk össze az adatokat

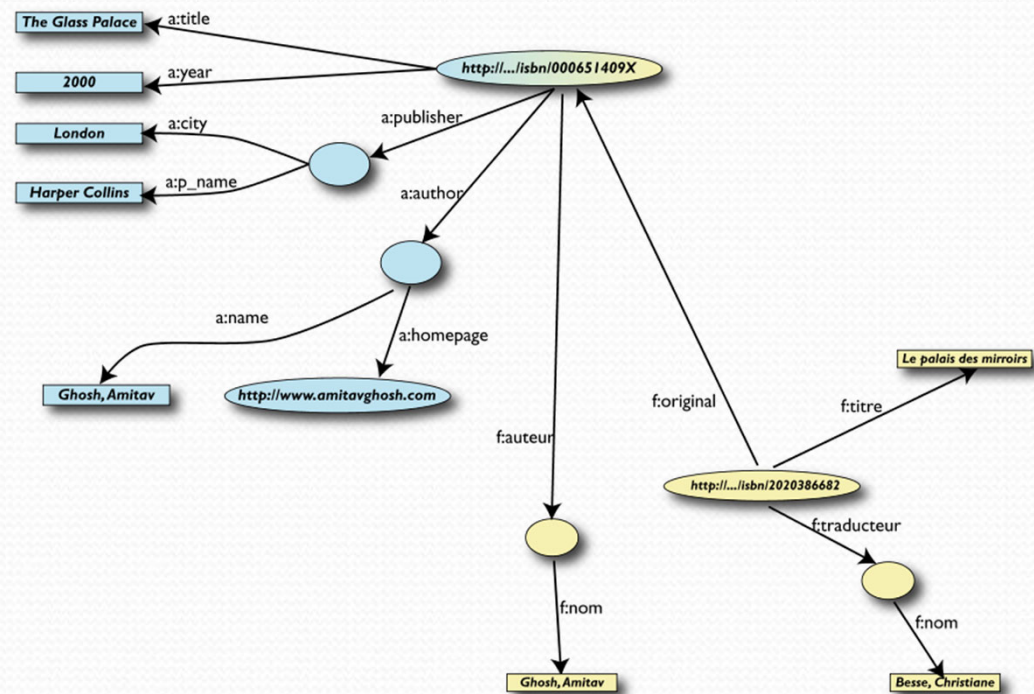


# Amennyiben identikusak az elemek:



# Írjunk lekérdezéseket a kapcsolatokon keresztül:

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

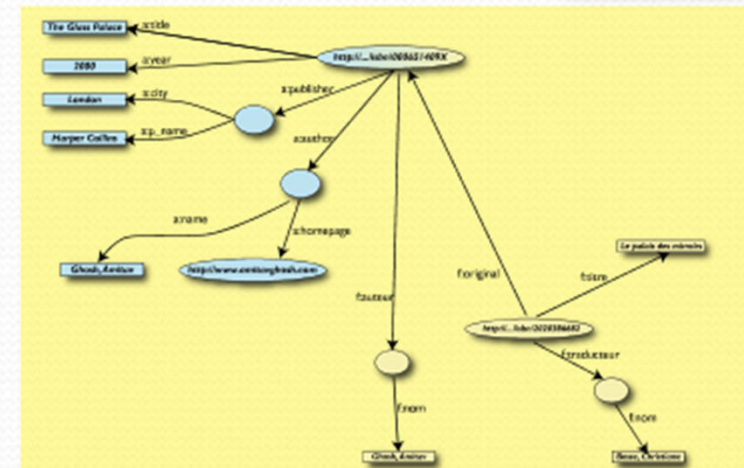
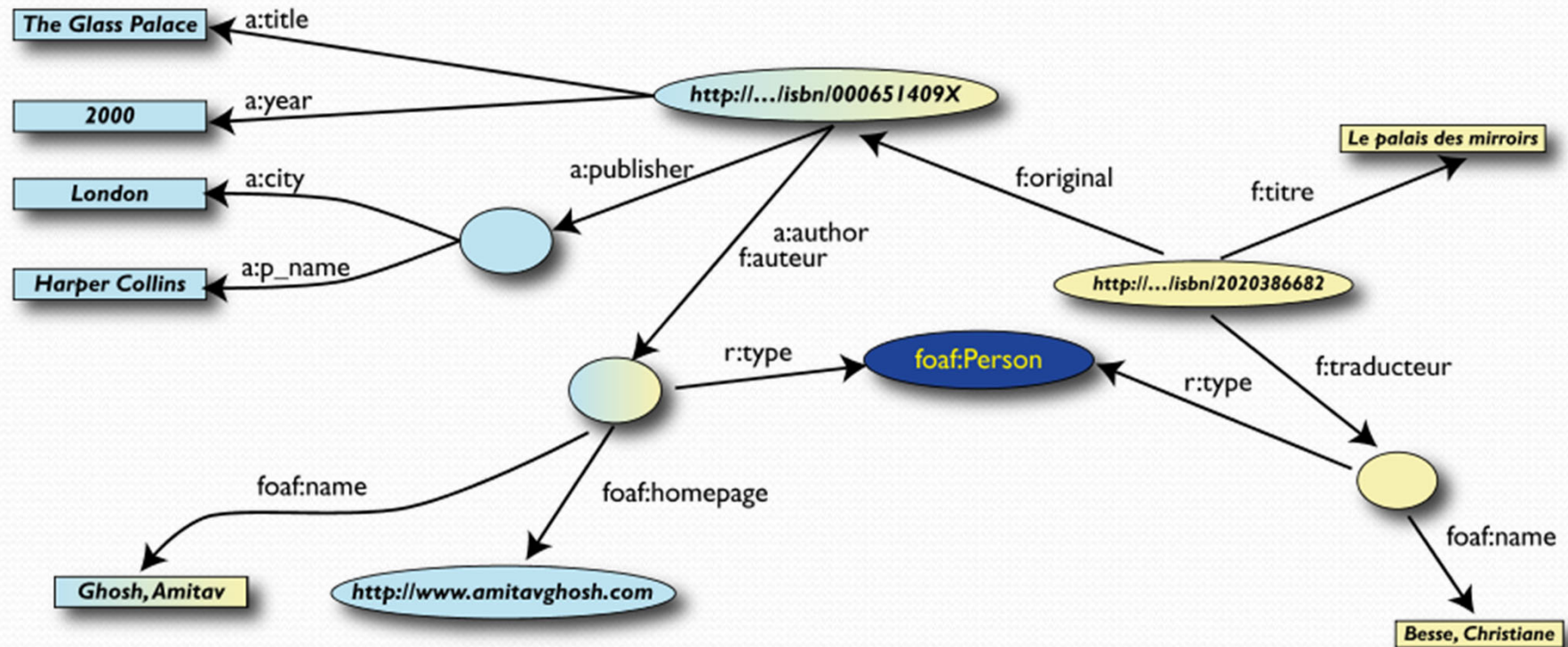


# További kapcsolatok is felfedezhetőek...

- 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



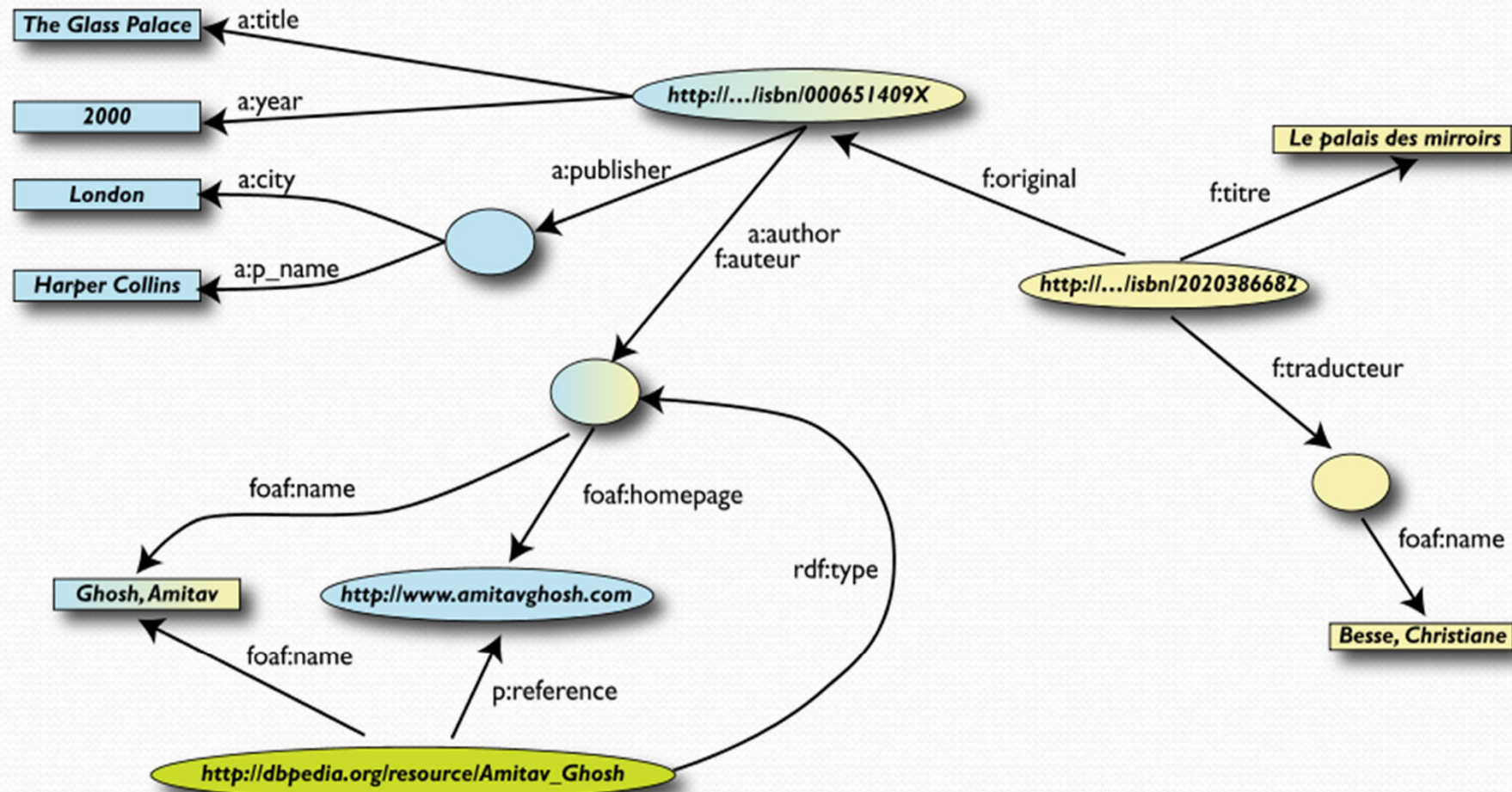
# Adatháló kiegészíthető, lekérdezhető így:



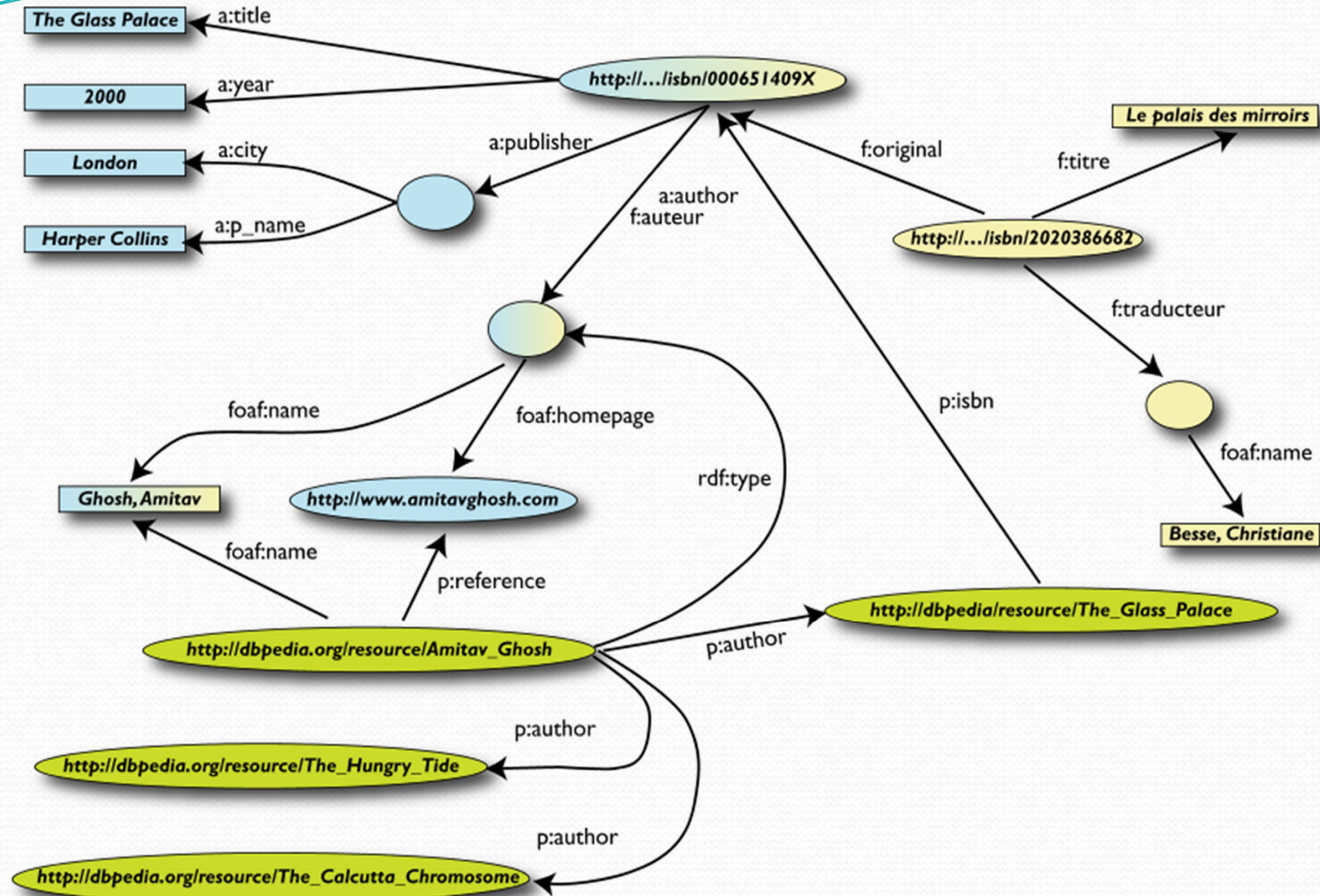
# Bővítés újabb adathalmazokkal

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

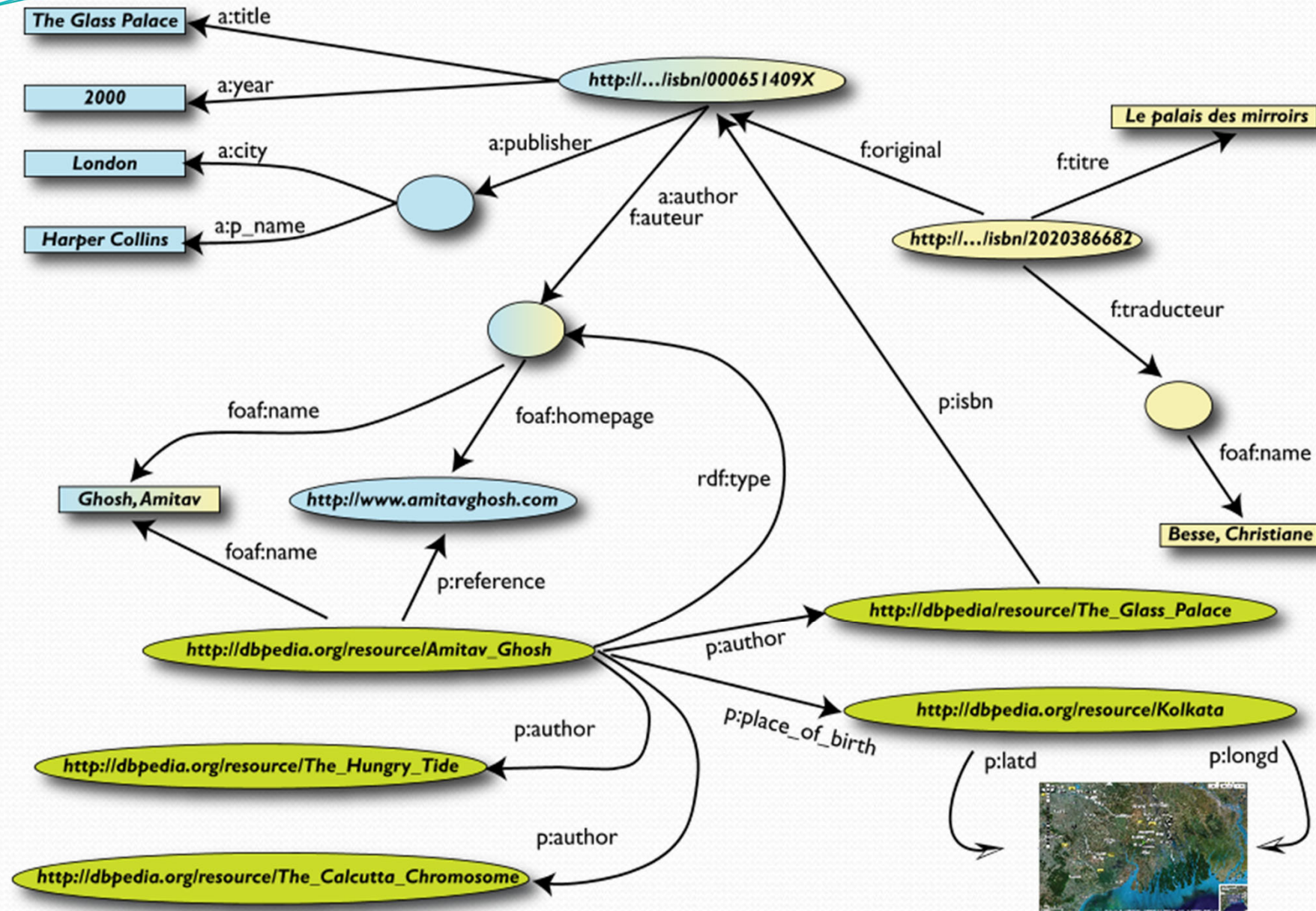
# Összekapcsolás Wikipedia adatokkal



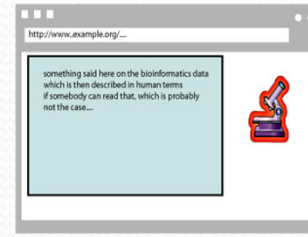
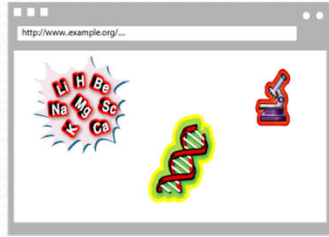
# Összekapcsolás Wikipedia adatokkal



# És így tovább...

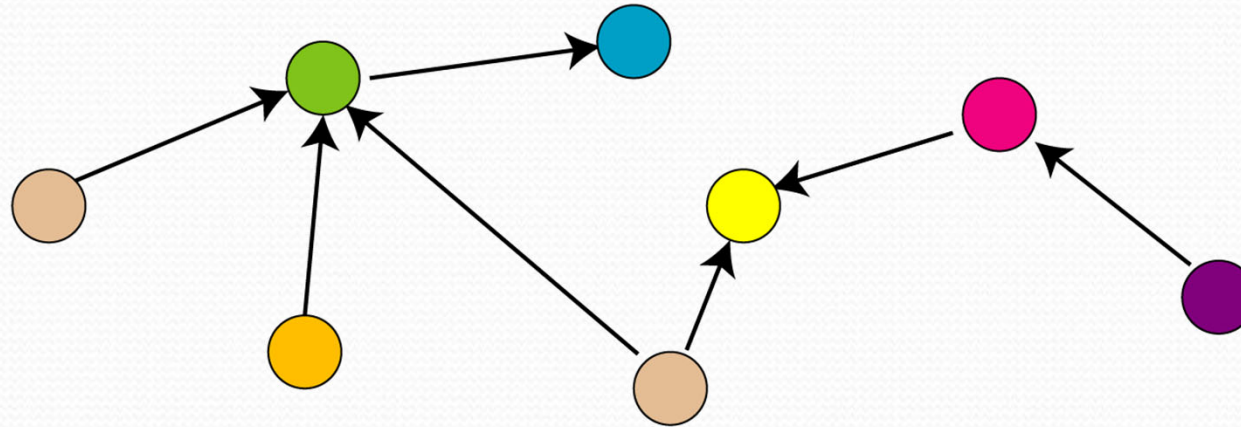


# Az adatok webje kialakulófélben...



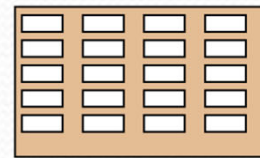
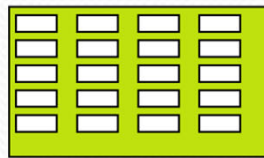
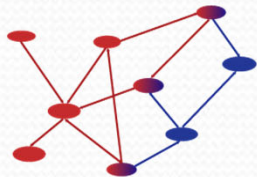
Alkalmazások

Lekérdezés,  
adatmódosítás



Absztrakt adatstruktúra

Leképezés,



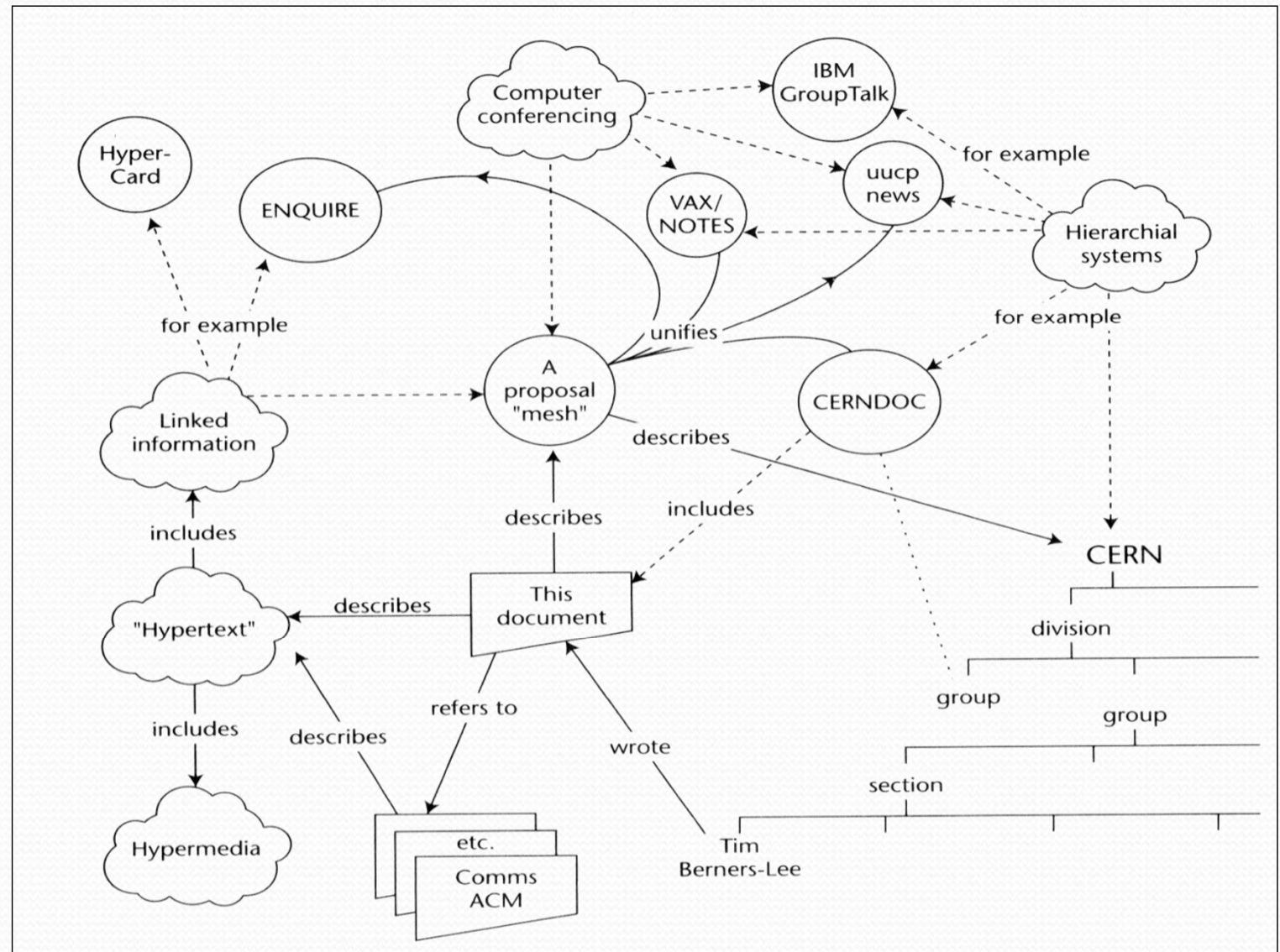
Forrás adatok különböző formátumokban

# Szemantikus Web

- 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.





# W3C szervezet célkitűzései

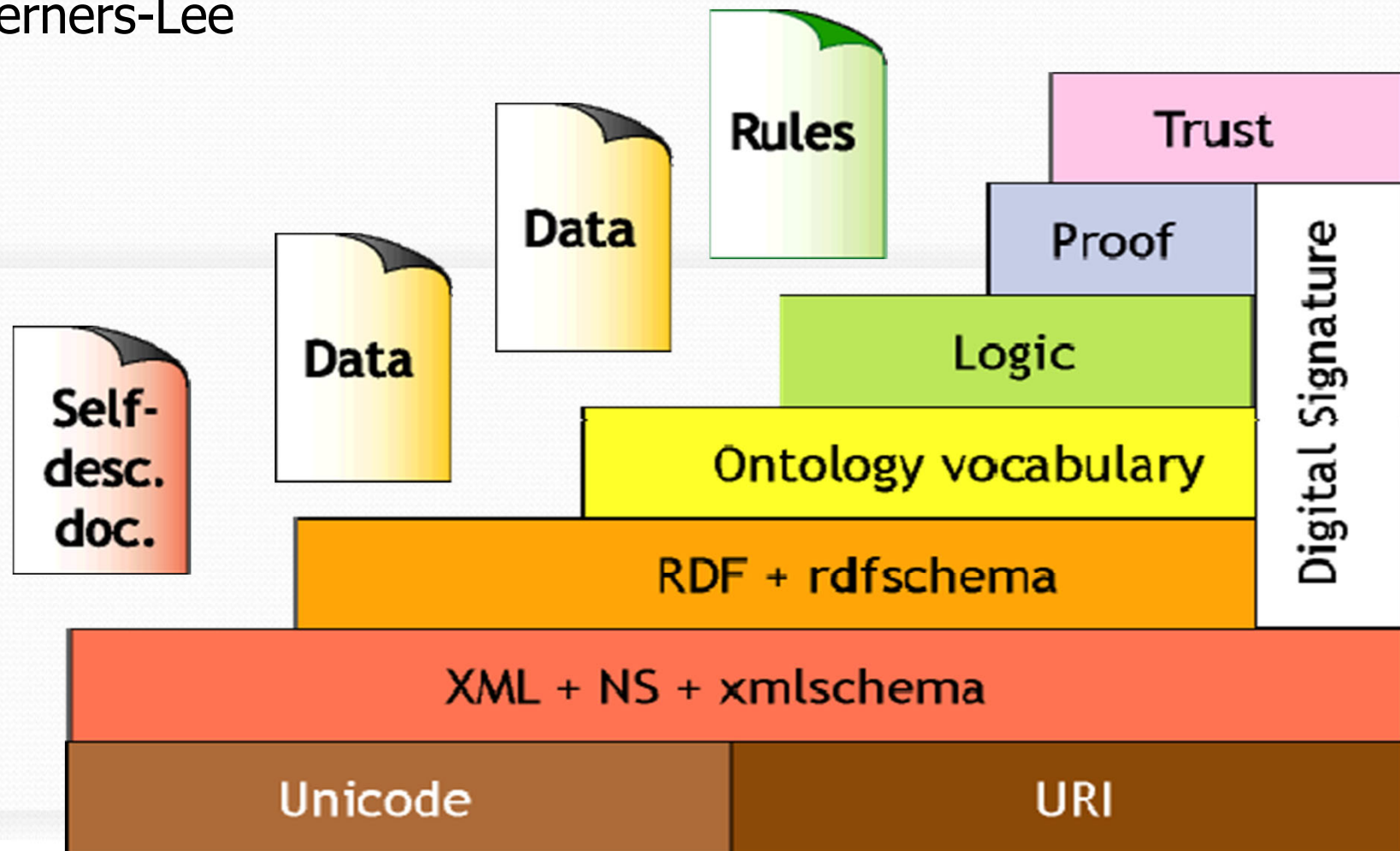
- Megközelítés – számítógépek jobb kihasználtságának biztosítása:

*„A szematikus 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



# Dokumentumok

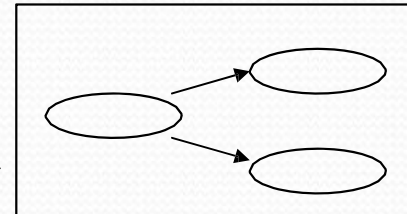
- **RDF Primer**
  - URI: <http://www.w3.org/TR/rdf-primer>
- **OWL Guide**
  - URI: <http://www.w3.org/TR/owl-guide/>
- **RDF Test Cases**
  - URI: <http://www.w3.org/TR/rdf-testcases/>
- **RDF: Concepts and Abstract Syntax**
  - URI: <http://www.w3.org/TR/rdf-concepts/>
- **RDF szemantika**
  - URI: <http://www.w3.org/TR/rdf-mt/>
  - Precíz, gráfokon alapuló szemantika
- **RDF/XML szintaxis**
  - URI: <http://www.w3.org/TR/rdf-syntax-grammar/>
- **RDF Vocabulary Description Language (RDF Schema)**
  - URI: <http://www.w3.org/TR/rdf-schema/>
- **Semantic Web/RDF Interest Group**
  - Vitafórum, alkalmazások
  - URI: <http://www.w3.org/RDF/Interest>
- **RDF Logic**
  - Nyilvános levelezési lista részletesebb szakmai vitákhoz
  - URI: <http://lists.w3.org/Archives/Public/www-rdf-logic/>
- **Annotation and Collaboration**
  - Nyilvános levelezési lista RDF-alapú annotációs rendszerekről
  - URI: <http://lists.w3.org/Archives/Public/www-annotation/>
- **W3C Semantic Web Home page**
  - URI: <http://www.w3.org/2001/sw/>

# RDF - az első SzW nyelv

## XML Kódolás

```
<rdf:RDF .....>  
  <...>  
  <...>  
</rdf:RDF>
```

## Gráf reprezentáció



**Jó a gépi felhasználás szempontjából**

**RDF  
Adat Modell**

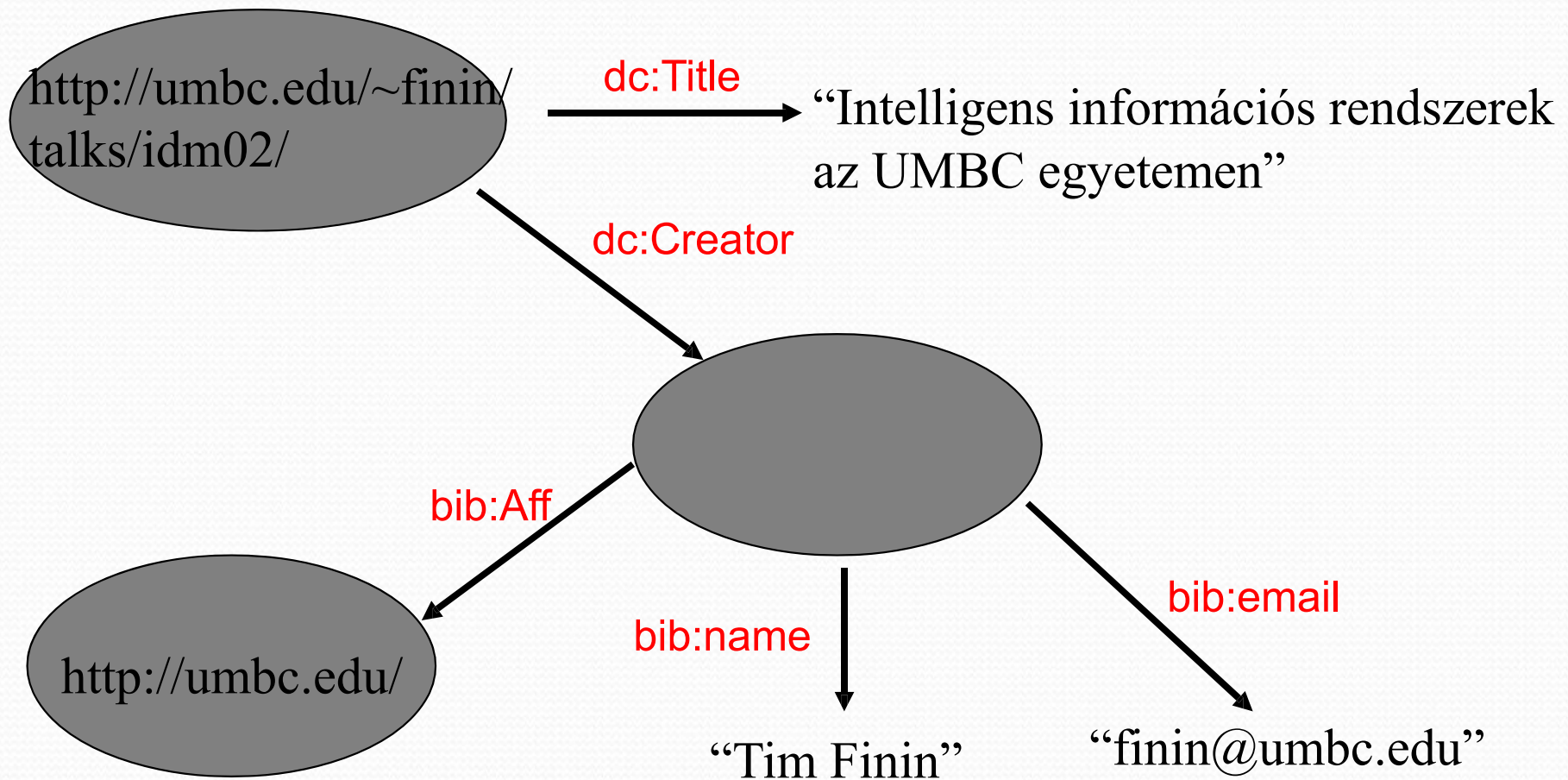
**Jól áttekinthető a felhasználónak**

## Predikátumok (hármások)

```
stmt(docInst, rdf_type, Document)  
stmt(personInst, rdf_type, Person)  
stmt(inroomInst, rdf_type, InRoom)  
stmt(personInst, holding, docInst)  
stmt(inroomInst, person, personInst)
```

**Alkalmas  
következtetésre**

# 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

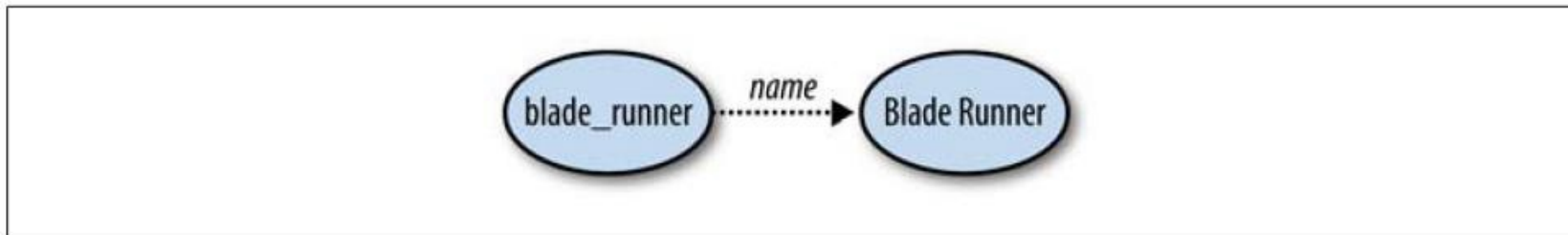
- Gráf adatmodell
- URI alapú szótárak
- Adattípusok
- Literálisok
- XML szerializációs szintaktika
- Egyszerű tények leírása
- Következtetés

# Graph adatmodell

- Hármások: alany, állítmány, tárgy
- Kifejezések: hármások gyűjteménye

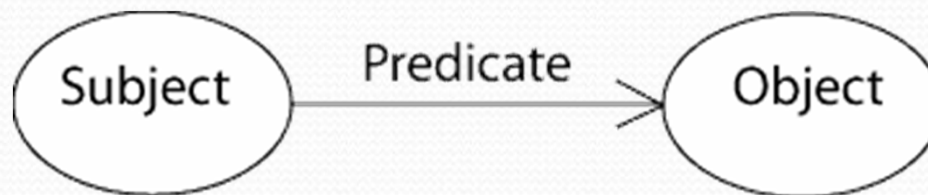


# Példa



# Terminológia és kényszerek

- Alany és tárgy a csomópontok
- Állítmány és tulajdonság szinonimák
- Különleges meg nem nevezett csomópontok: üres csomópontok
- Alany: URI referencia vagy üres csomópont
- Állítmány: URI referencia
- Tárgy: URI referencia, literális, üres csomópont



# Információ a hármásokban

<http://xmlns.com/foaf/0.1/workplaceHomepage>

<http://directory.com/people#FulvioCorno>

<http://www.polito.it/>

**RDF**

CompanyHomePage

PersonID	Homepage
FulvioCorno	<a href="http://www.polito.it/">http://www.polito.it/</a>

**Relációs adatbázis**

**Első rendű logikai  
predikátum**

```
HasCompanyHomePage(  
  'FulvioCorno',  
  'http://www.polito.it/');
```

# Hármasok vs adatbázisok

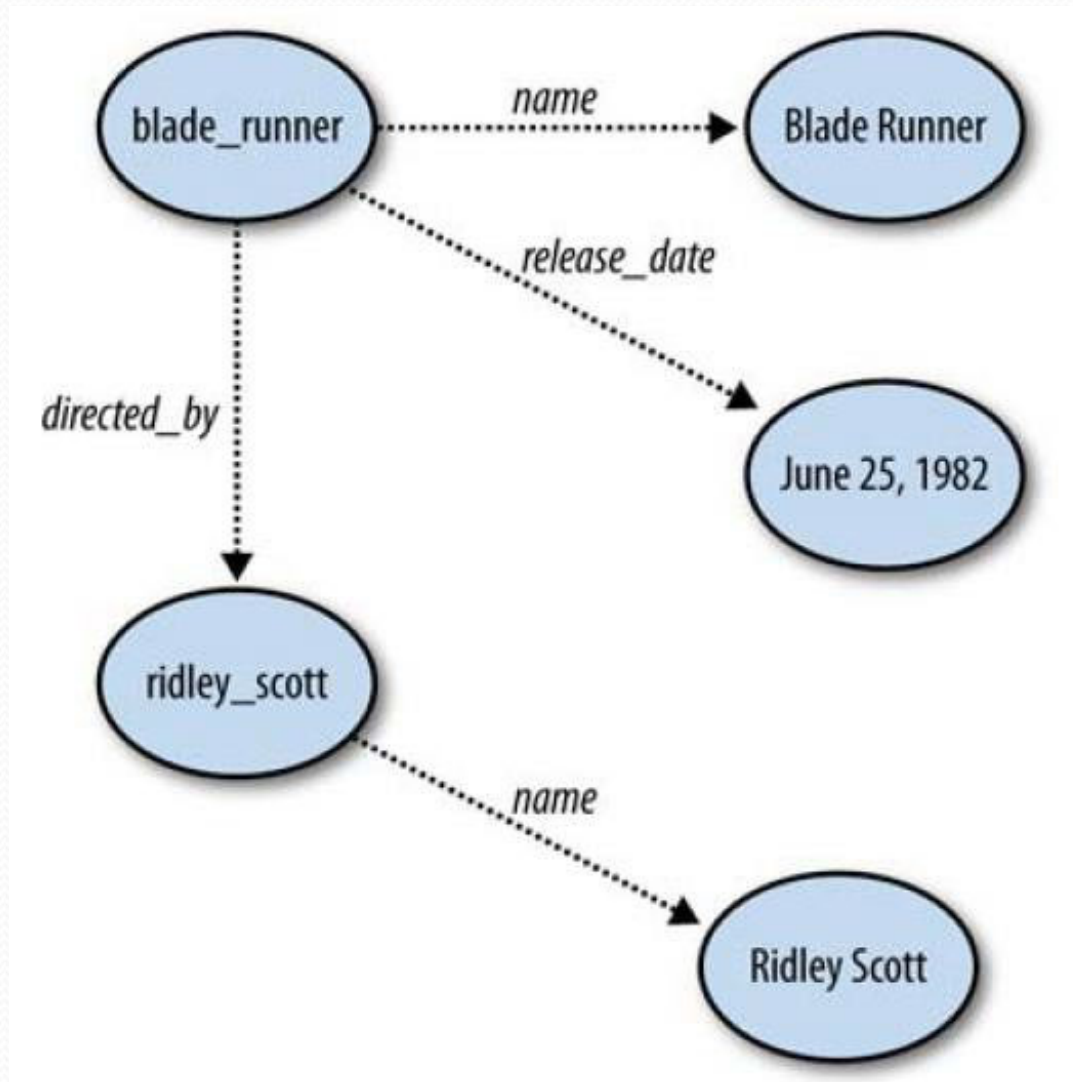
<b>FEATURE</b>	<b>RELATIONAL DATABASE</b>	<b>KNOWLEDGEBASE</b>
Structure	Schema	Ontology statements
Data	Rows	Instance statements
Administration language	DDL	Ontology statements
Query language	SQL	SPARQL
Relationships	Foreign keys	Multidimensional
Logic	External of database/triggers	Formal logic statements
Uniqueness	Key for table	URI



# Összehasonlítva...

- Relációs adatbázisban tetszőleges számú oszlopot definiálhatunk
- Elsőrendű logikában a predikátumok tetszőleges elemet (argumentumot) tartalmazhatnak
- Egy RDF hármass csak egy alanyt és egy tárgyat tartalmazhat
  - Komplex kijelentéseket dekomponálni kell

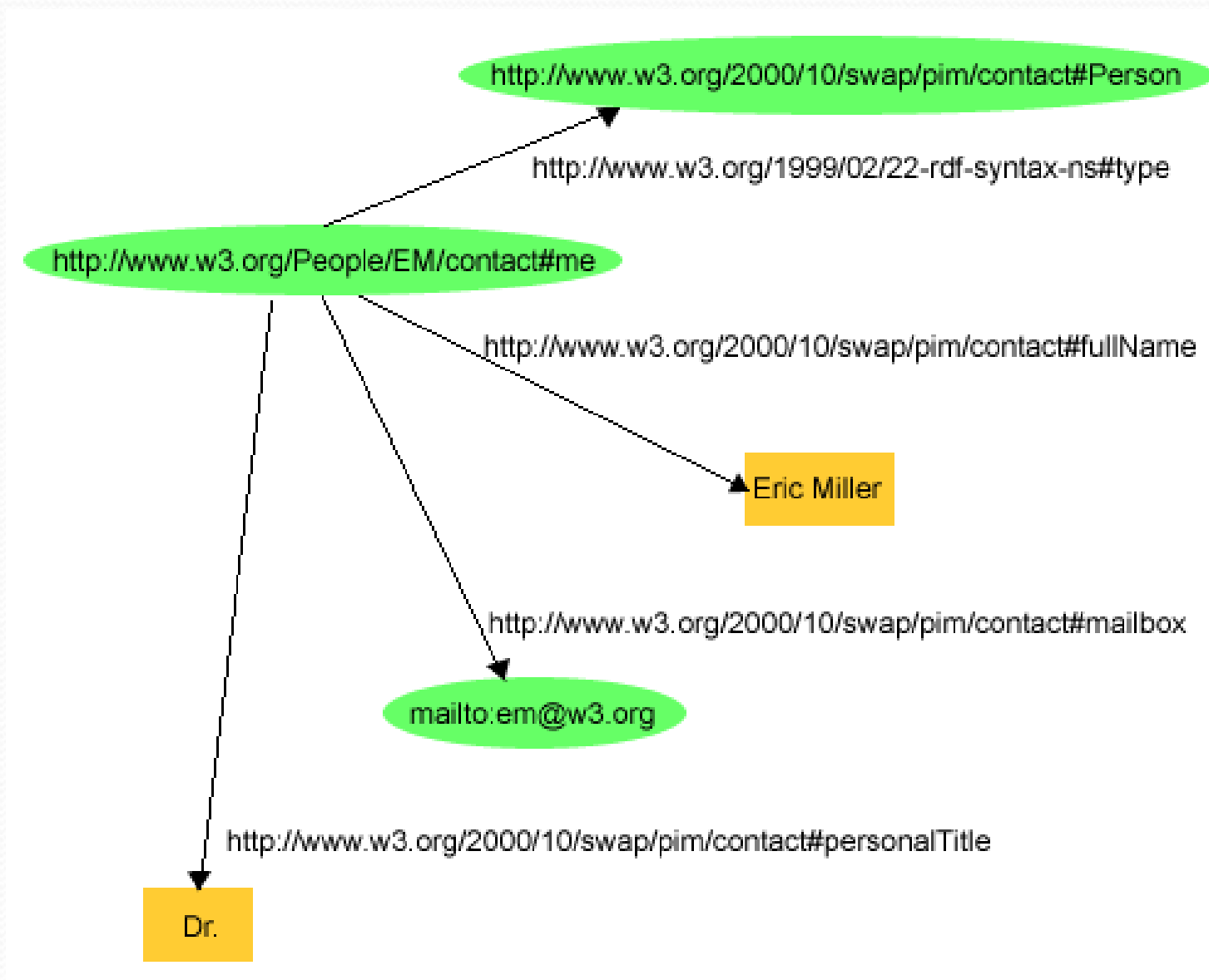
# Példa



# Példa

- Írjuk le RDF állításokkal a következő kijelentést
- "létezik egy személy, akit a <http://www.w3.org/People/EM/contact#me> oldallal azonosíthatunk, akinek neve Eric Miller, email címe [em@w3.org](mailto:em@w3.org), és a titulusa Dr."

# Példa



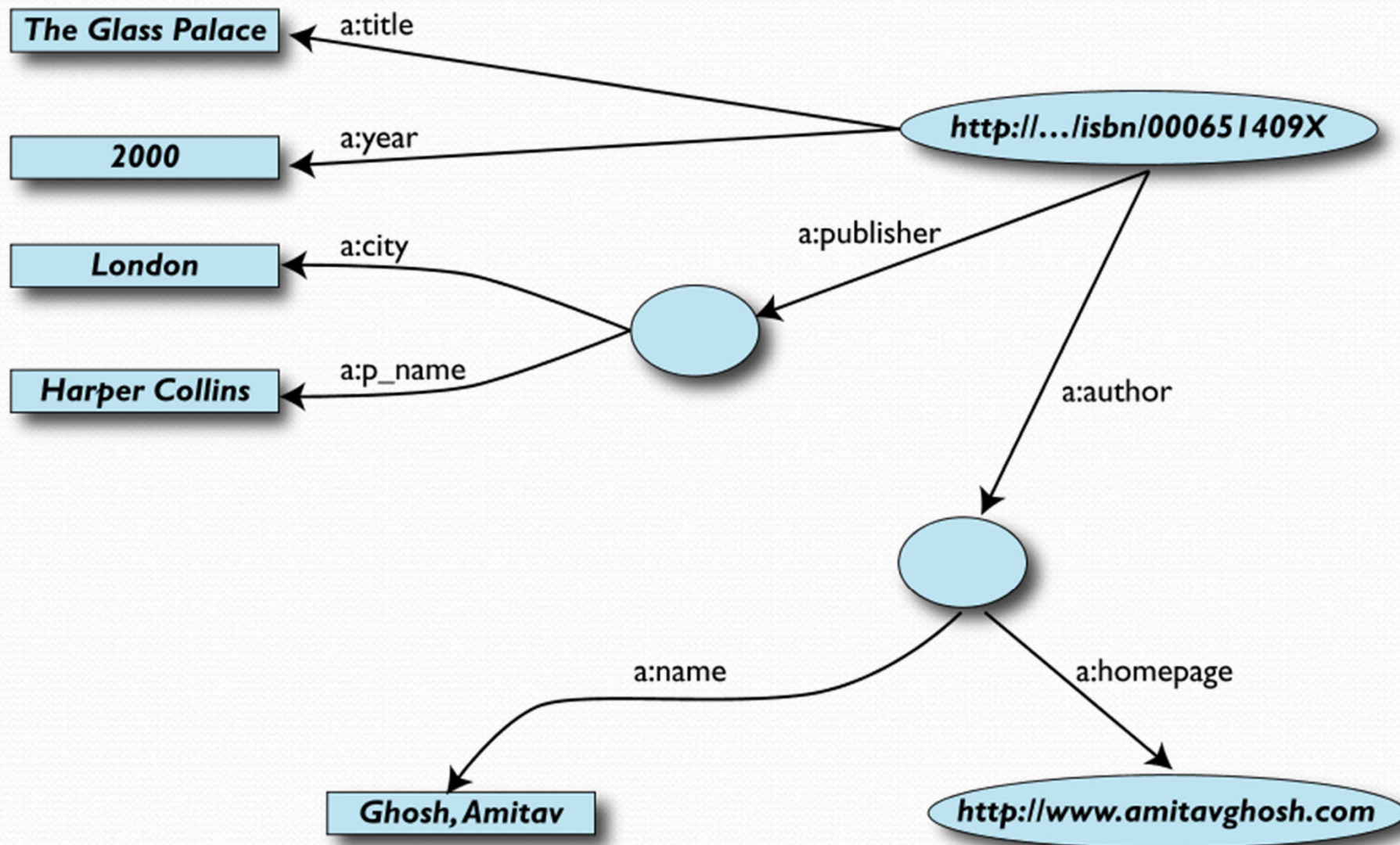
# Példa: könyvesbolt adatai

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

<b>ID</b>	<b>Name</b>	<b>Home Page</b>
id_xyz	Ghosh, Amitav	<a href="http://www.amitavghosh.com">http://www.amitavghosh.com</a>

<b>ID</b>	<b>Publ. Name</b>	<b>City</b>
id_qpr	Harper Collins	London

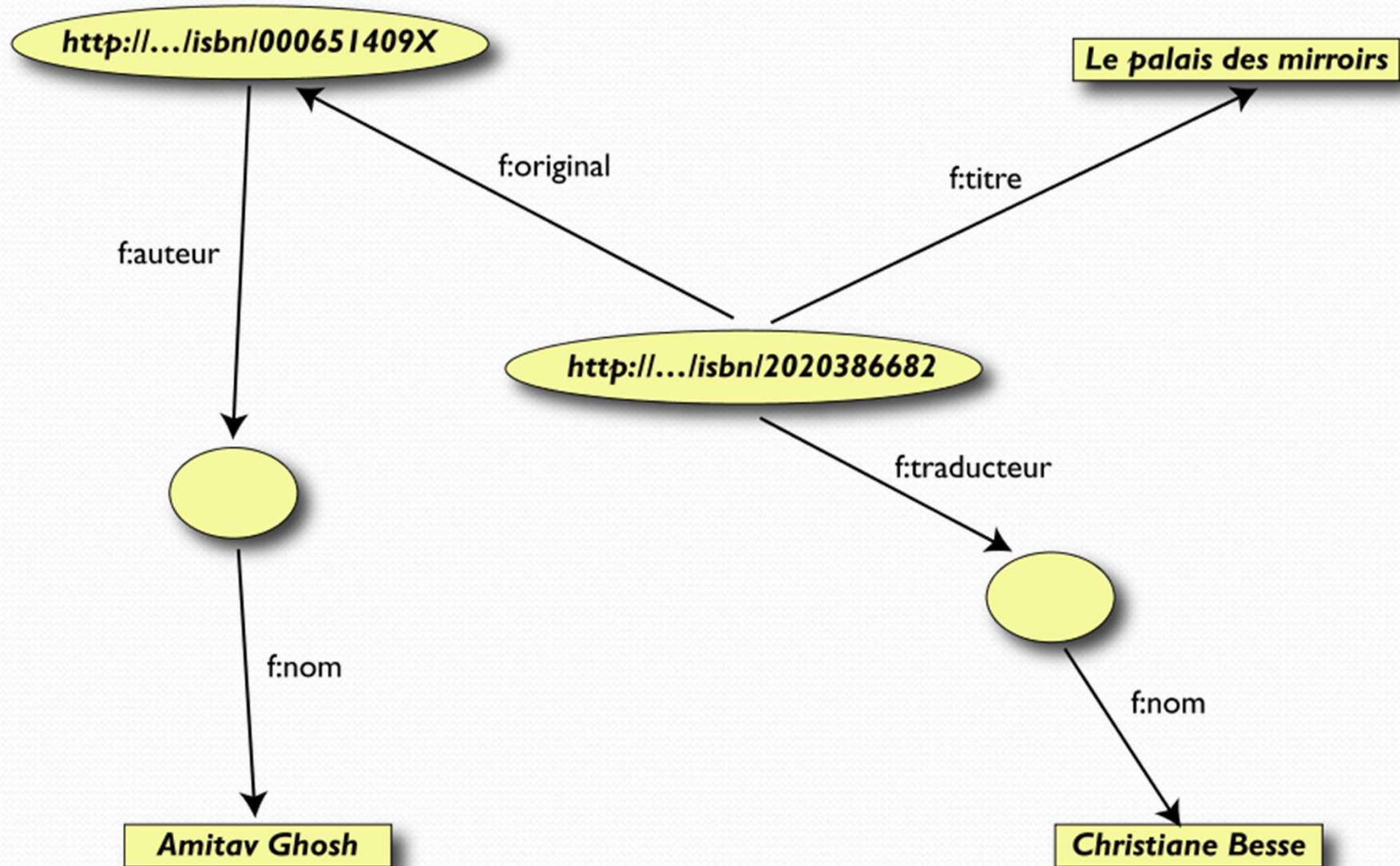
# Adatok exportálása relációkként



# Egy másik könyvesbolt adatai

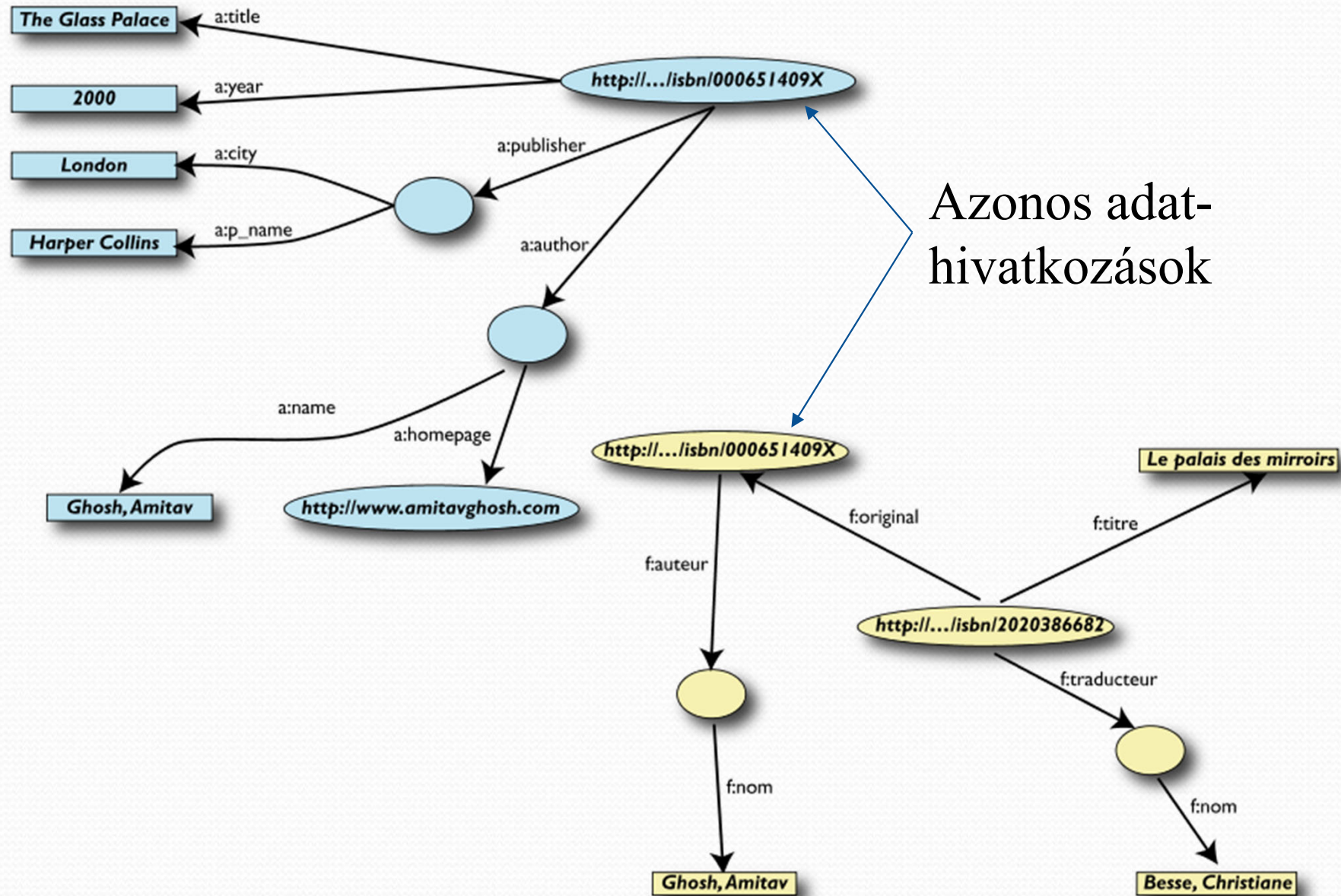
	A	B	D	E
1	<b>ID</b>	<b>Titre</b>	<b>Traducteur</b>	<b>Original</b>
2	ISBN0 2020386682	Le Palais des miroirs	A13	ISBN-0-00-651409-X
3				
6	<b>ID</b>	<b>Auteur</b>		
7	ISBN-0-00-651409-X	A12		
11	<b>Nom</b>			
12	Ghosh, Amitav			
13	Besse, Christianne			

# A második könyvesbolt adatainak exportja

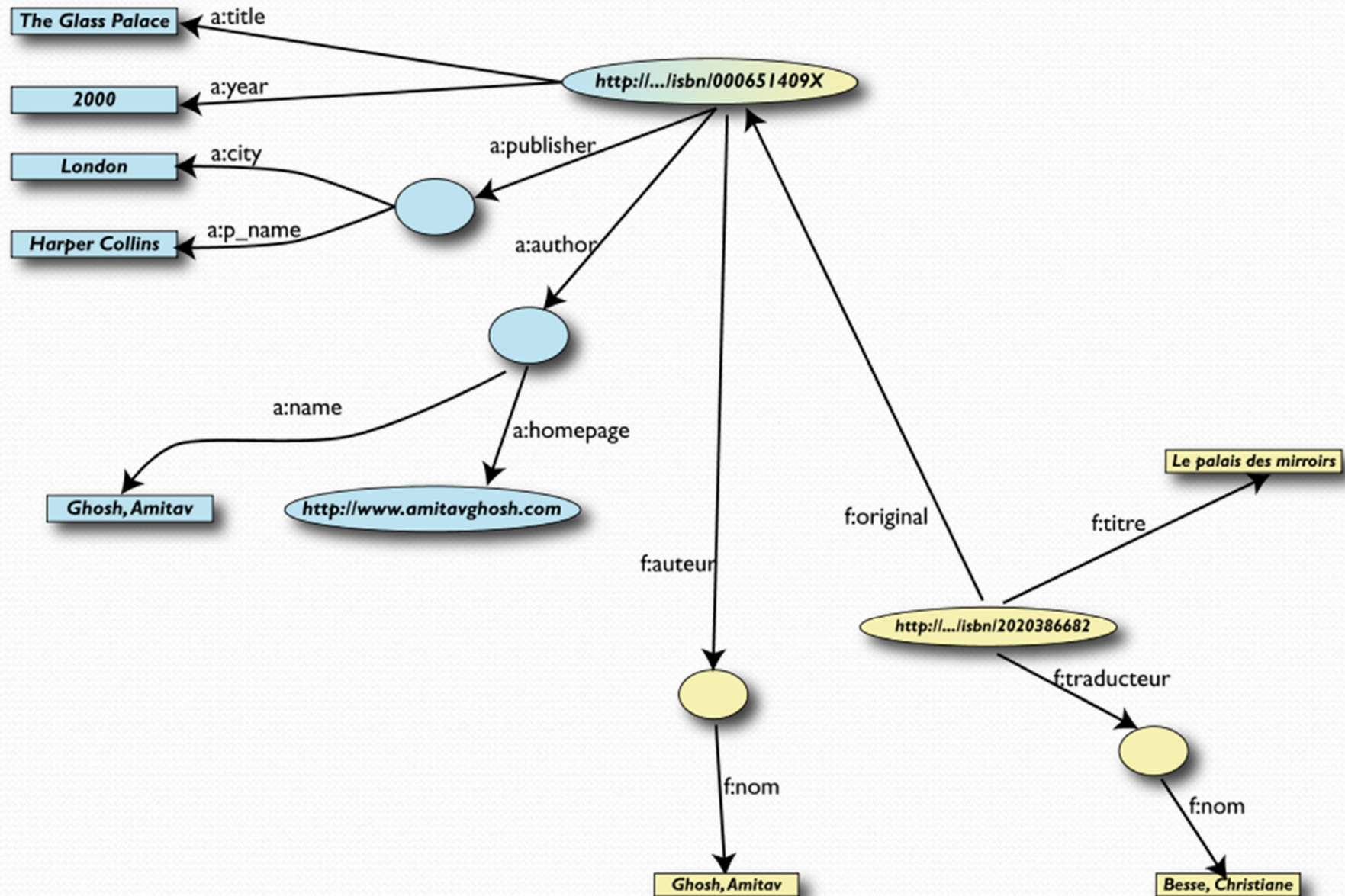




# Kapcsoljuk össze az adatokat

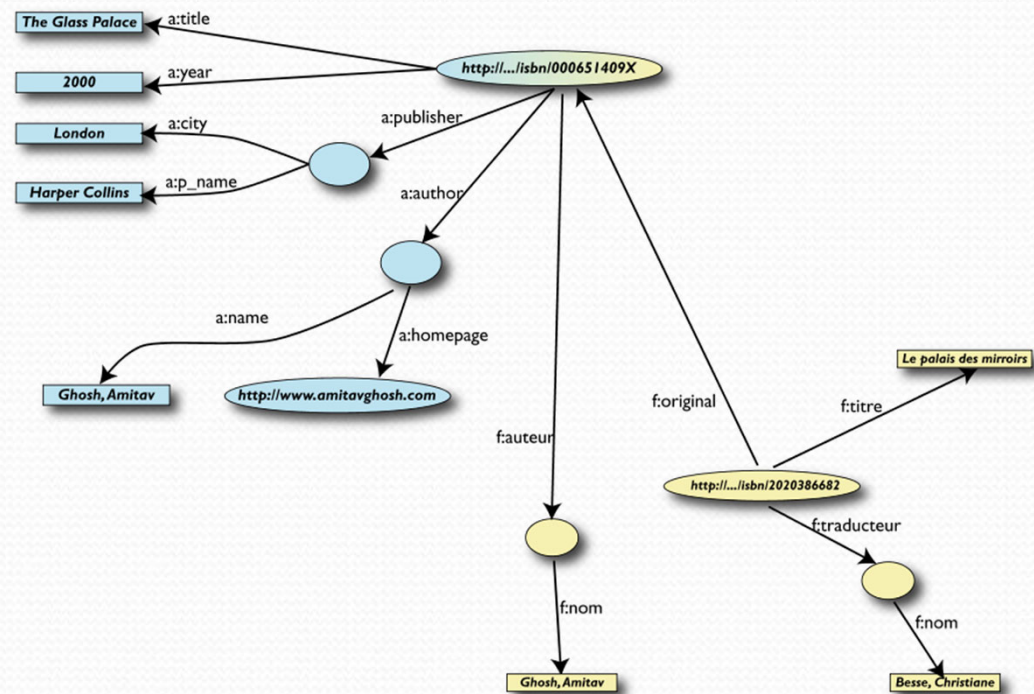


# Amennyiben identikusak az elemek:



# Írjunk lekérdezéseket a kapcsolatokon keresztül:

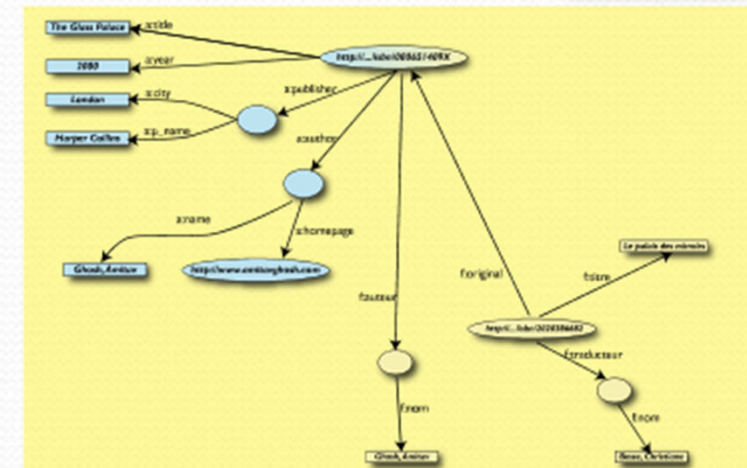
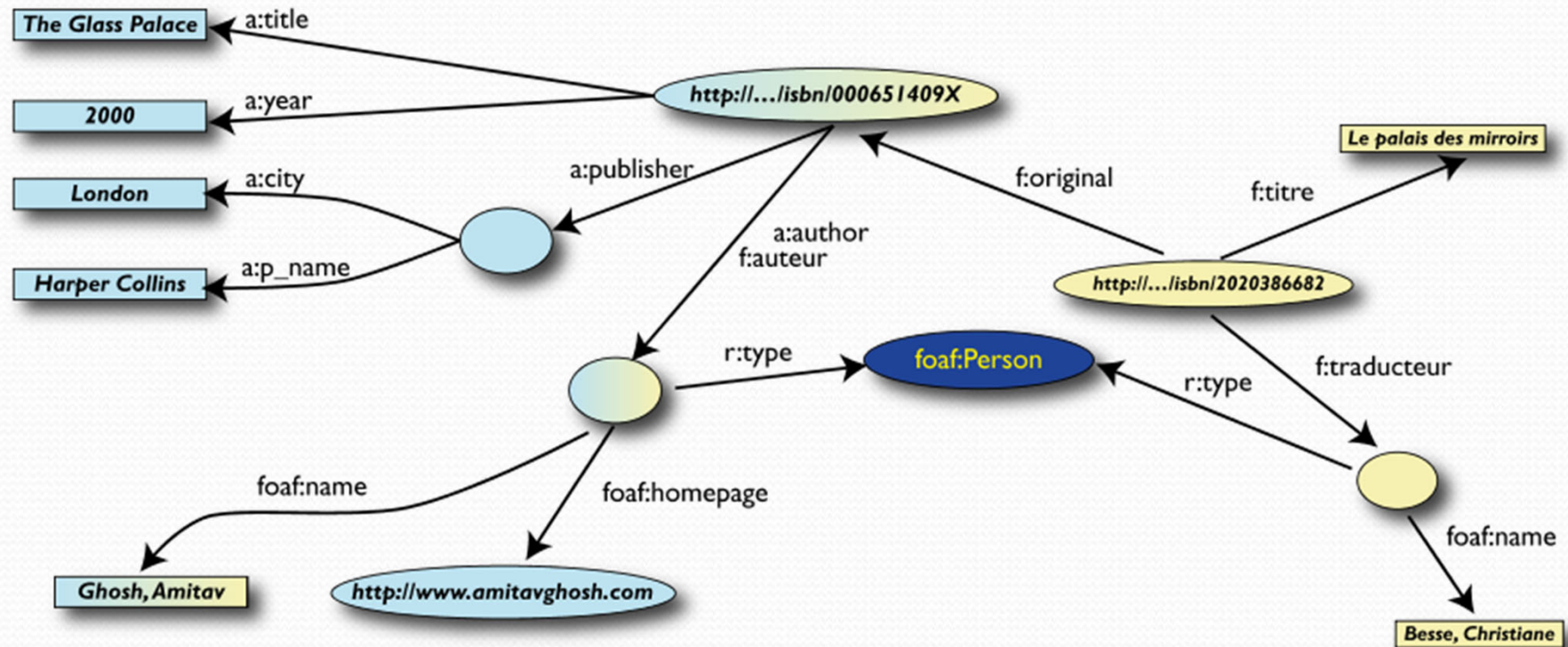
- Az első könyvesbolt adatai kiegészíthető például az eredeti könyvre vonatkozó információkkal



# 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:



# RDF áttekintés

## **A formális modell alapelemei:**

Két alaphalmaz: erőforrások (resources) és literálisok (literals)

Az erőforrások egy fontos részhalmaza: Tulajdonságok (properties).

Definiálunk egy hármassokból álló halmazt: Állítások (statements), amelyek formája: {alany, állítmány, tárgy}

, ahol az alany egy erőforrás,  
az állítmány egy tulajdonság,  
a tárgy vagy erőforrás vagy literális.

# RDF adatmodell

- Erőforrások (Resources)
  - URI azonosítja
  - Kijelentés vonatkozik rá
- Tulajdonságok (Properties)
  - Erőforráshoz kapcsolt jellemző
  - A tulajdonság is erőforrás
- Literálok (Literals)
  - Karakter sorozatok
- Kijelentések (Statements)
  - Alany (subject); erőforrás
  - Állítmány (predicate); tulajdonság
  - Tárgy (object); erőforrás vagy literál

## Ezek segítségével megfogalmazhatunk leírásokat:

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:s="http://description.org/schema/">
  <rdf:Description about="http://www.w3.org/Home/Lassila">
    <s:Creator>
      <rdf:Description about="http://www.w3.org/staffId/85740">
        <rdf:type resource="http://description.org/schema/Person"/>
        <v:Name>Ora Lassila</v:Name>
        <v:Email>lassila@w3.org</v:Email>
      </rdf:Description>
    </s:Creator>
  </rdf:Description>
</rdf:RDF>
```



# Az RDF szerepe az SZW hierarchiában

- Technológiát és módszert ad ahhoz, hogy dokumentumainkhoz jelentést rendeljünk egy jól olvasható formában
- Jó lehetőség, de
  - a szemantikai információk nem túl hasznosak, amíg strukturálatlan és nem tudjuk konzisztens módon értelmezni.

(XML séma kevés: csak szintaktikáról szól, nem ad lehetőséget a dokumentumon kívüli dolgok leírására)

# Séma hiányában ugyanannak a tartalomnak sokféle reprezentációja lehetséges:

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-
    syntax-ns#" />
  <rdf:Description
    about="http://www.w3.org/Home/Lassila">
    <Creator>
      <rdf:Description
        about="http://www.w3.org/staffId/85740">
        <rdf:type
          resource="http://desc.org/schema/Person"/>
        <Name>Ora Lassila</Name>
        <Email>lassila@w3.org<Email>
      </rdf:Description>
    </Creator>
  </rdf:Description>
</rdf:RDF>
```

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-
    syntax-ns#" />
  <rdf:Description
    about="http://www.w3.org/Home/Lassila">
    <author>
      <rdf:Description
        about="http://www.w3.org/staffId/85740">
        <rdf:type
          resource="http://desc.org/schema/Person"/>
        <name>
          <surname>Lassila</surname>
          <given>Ora</given>
        </name>
        <email>lassila@w3.org</email>
      </rdf:Description>
    </author>
  </rdf:Description>
</rdf:RDF>
```

# SPARQL

- Lekérdező nyelv RDF-hez
- SPARQL gráf illesztésen alapuló lekérdező nyelv
- Gráf minták

- példa:

```
<http://example.org/book/book1>  
<http://purl.org/dc/elements/1.1/title>  
?title .
```

- ?title - változó.

# Egy egyszerű SPARQL Query

- **Adat:**

```
<http://example.org/book/book1>  
  <http://purl.org/dc/elements/1.1/title>  
  "SPARQL Tutorial".
```

- **Lekérdezés:**

```
SELECT ?title  
WHERE { <http://example.org/book/book1>  
        <http://purl.org/dc/elements/1.1/title>  
        ?title . }
```

- **Eredmény:**

<b>title</b>
"SPARQL Tutorial"

# További példa

- **Adat:**

```
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
```

```
_ :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:mbox <mailto:carol@example.org> .
```

- **Lekérdezés:**

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

- **Eredmény:**

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

# Lekérdezések RDF literálisokkal

- **Példa RDF adatokra**

```
@prefix dt: <http://example.org/datatype#> .
@prefix ns: <http://example.org/ns#> .
@prefix : <http://example.org/ns#> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .

:x ns:p "cat"@en .
:y ns:p "42"^^xsd:integer .
:z ns:p "abc"^^dt:specialDatatype .
```

# RDF Literálisok illesztése

- Lekérdezés 1:

```
SELECT ?v WHERE { ?v ?p "cat" }
```

## Lekérdezés 2:

```
SELECT ?v WHERE { ?v ?p "cat"@en }
```

eltérő eredményt ad.

- Csak a második találja meg az előző példában az eredményt:

<b>v</b>
<code>&lt;http://example.org/ns#x&gt;</code>



# Üres csomópontok a lekérdezésekben

- **Adat:**

```
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
```

```
_:a foaf:name "Alice" .
```

```
_:b foaf:name "Bob" .
```

- **Lekérdezés:**

```
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
```

```
SELECT ?x ?name
```

```
WHERE { ?x foaf:name ?name . }
```

- **Eredmény:**

<b>x</b>	<b>name</b>
_:c	"Alice"
_:d	"Bob"

# Üres csomópontok a lekérdezésekben (modell bővítés)

- **Adat:**

```
@prefix foaf: <http://xmlns.com/foaf/0.1/> .  
  
_:a foaf:name "Alice" .  
_:b foaf:name "Bob" .  
_:a foaf:knows _:b .  
_:b foaf:knows _:a .
```

- **Lekérdezés:**

```
PREFIX foaf: <http://xmlns.com/foaf/0.1/>  
SELECT ?x ?name1 ?y ?name2  
WHERE { ?x foaf:name ?name1 . ?y foaf:name ?name2 .  
        ?x foaf:knows ?y }
```

- **Eredmény:**

?x	name1	?y	name2
_:c	"Alice"	_:d	"Bob"
_:d	"Bob"	_:c	"Alice"

# RDF esettanulmányok

- Dublin Core
  - Magas szintű szótár definiálása
  - Elektronikus dokumentumok megtalálása
- Open Directory Project (OPD)
  - Webes katalógus keresők számára
- MusicBrainz
  - Hanganyagok (cd, mp3 ...) metaadatainak lekérésére
- RSS: RDF Site Summary
  - Hírek, események közzététele
- Wordnet
  - Szabadon letölthető szótár
  - Nem csak címszavakat, hanem kapcsoltaokat is leír

# 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	...	Fn
M1	E1,1	E1,2			
M2		E2,2			
M3			E3,3		
...				...	
Mm					Em,n

Többértelmű szavak

# Reprezentációs módszer

## - Konstruktív

- A reprezentáció tartalmazzon elégséges információt a fogalom felépítéséhez

## - Differenciális

- A jelentések úgy legyenek fűzésekkel reprezentálva, hogy megkülönböztethetőek 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

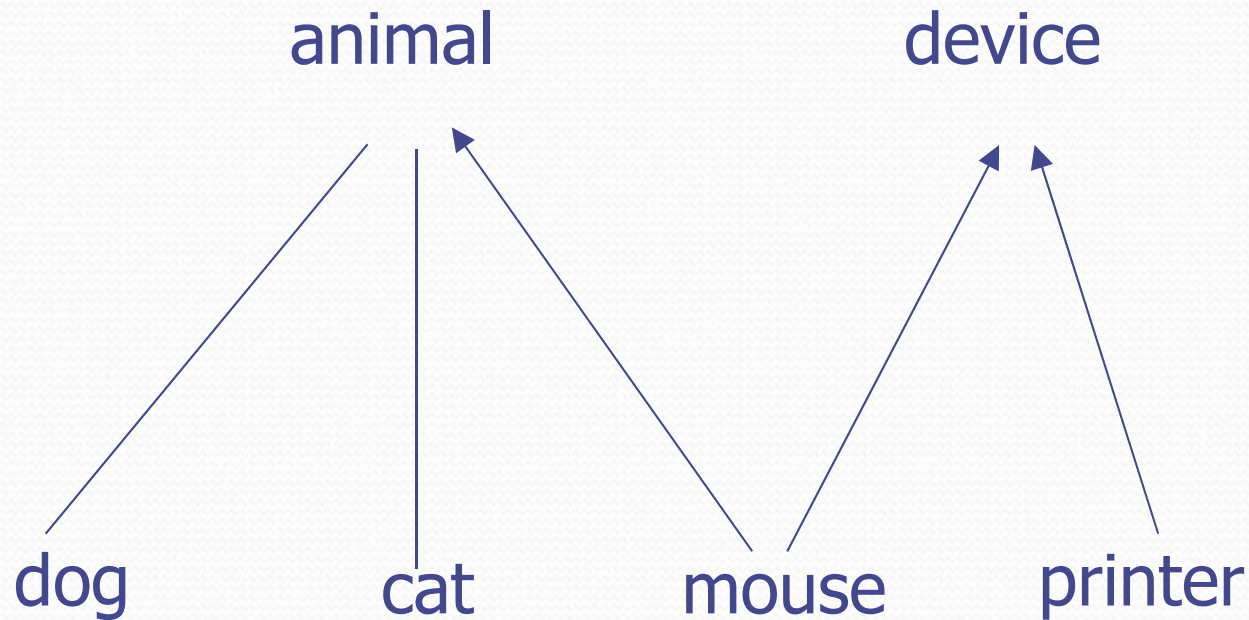
Két szó szinonima, azonos az értelmük, ha a következő kapcsolatok fennállnak:

- 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 fogalom jelentését, de megjeleníti, igazolja a fogalom létezését*

# Hiponima

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



# Meronymia/Holonymia

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