

# *Introduction to the Semantic Web (tutorial)*

***2009 Semantic Technology Conference  
San Jose, California, USA  
June 15, 2009***

*Ivan Herman, W3C  
[ivan@w3.org](mailto:ivan@w3.org)*



# *Introduction*

*Let's organize a trip to Budapest using the Web!*

*You try to find a proper flight with ...*

# ... a big, reputable airline, or ...

Book flights - Mozilla Firefox

File Edit View History Bookmarks Tools Help

[http://www.klm.com/travel/nl\\_nl/apps/ebt/ebt\\_home.htm](http://www.klm.com/travel/nl_nl/apps/ebt/ebt_home.htm)

KLM Royal Dutch Airlines

nwa

Stel hier uw vraag (in me) ?

You are not logged in

Contact KLM Service Center

Book flights

- 1 Vlucht zoeken
- 2 Datum kiezen
- 3 Tijd kiezen
- 4 Uw gegevens
- 5 Bekijken & betalen

E-mail een vriend(in) Print deze pagina

Kies uw heen- en terugvlucht.

Vertrek: Amsterdam (Schiphol) naar Boedapest (Ferihegy Airport)

Alleen vluchten gebaseerd op Alleen de beste verbindingen (10 van 13)  
 Alleen rechtstreekse vluchten tonen (5 van 13)  
 Laagste tarieven (9 van 13)  
 Alle vluchten (13)

Kies	Prijs	Vertrek	Aankomst	Vlucht
<input checked="" type="radio"/> Rechtstreekse vluchten		10:00 Vr 15 Aug 08 Amsterdam (Schiphol) Totale reistijd: 2 uren 0 minuten Uitgevoerd door MALEV Hungarian Airlines Vliegtuigtype: Boeing 737	12:00 Vr 15 Aug 08 Boedapest (Ferihegy Airport)	KL3201
<input type="radio"/>		12:20 Vr 15 Aug 08 Amsterdam (Schiphol) Totale reistijd: 2 uren 0 minuten Uitgevoerd door MALEV Hungarian Airlines Vliegtuigtype: Boeing 737	14:20 Vr 15 Aug 08 Boedapest (Ferihegy Airport)	KL3203

**Kies tarief**

**Take Off**

RETOUR tarieven per persoon incl. belasting en toeslagen (excl. reserveringskosten)

Take Off (25% FB Miles)	300
<input checked="" type="checkbox"/> Geselecteerd	
Take Off	341
<input type="checkbox"/> Selecteer	
Take Off (Flexibel)	383
	455
	537

# *... the airline of the target country, or ...*

Review - Mozilla Firefox

File Edit View History Bookmarks Tools Help

<https://booking.malev.com/BookingSite/Review/Review.aspx>

BOOK YOUR TRIP  
FLIGHTS, SCHEDULES  
PREPARE FOR FLIGHT  
FLYING WITH MALEV  
DUNA CLUB  
CUSTOMER CARE  
CORPORATE INFO

MALEV Hungarian Airlines ONE MORE REASON TO TRAVEL

English

1 Search 2 Select 3 Review 4 Traveller data 5 Purchase 6 Confirmation

SELECT DELIVERY METHOD

Please select ticket delivery method below!

Delivery method	Service fee	Address
<input checked="" type="radio"/>  E-ticket	EUR 10 <a href="#">View service fee breakdown</a>	with e-invoice
<input type="radio"/>  Airport	EUR 30	KLM ticket office, departure hall 2.
<input type="radio"/>  Office	EUR 30	KLM ticket office, departure hall 2.
<input type="radio"/>  Courier		Courier or mail delivery is available only within the country of departure! You will be asked to provide a delivery address on the next page. Please note that in case of courier delivery we will be unable to deliver your ticket(s) to a PO box.
<input type="radio"/>  Mail	EUR 30	

BOOKING GUIDE

Book Cheap!  
Card payment  
About ticket pickup  
If you are not among the travellers  
Online Client Service

  
Check it!

FLIGHT SUMMARY

Outbound flight	Return flight
From: Amsterdam, Schiphol (AMS), Netherlands Budapest, Ferihegy 2A (BUD).	From: Budapest, Ferihegy 2A (BUD), Hungary To: Amsterdam, Schiphol (AMS),

# ... or a low cost one

Online booking | Select Flights - Mozilla Firefox

File Edit View History Bookmarks Tools Help

<http://wizzair.com/skylights/cgi-bin/skylights.cgi?step=1>

wizzair.com

new destination from London Luton:  
Timisoara (from 17 December)

WIZZ

online booking useful information destinations travel services partners

tant changes regarding checked-in baggages - Please click here for more details...

search  select  contact  passengers  purchase  confirmation

The flights available for the date(s) that you have selected are shown below. Review and select that you wish to purchase by ticking the dot next to the fare price or use the form to the left hand side to search for new flights. All times are local.

Fares shown below are for one way flights and per adult, child and infant. The total price includes the taxes and the charges. Payments made with debit and credit cards are subject to a payment fee. Click here to find out the exact amount. The fee depends on the type of card that you wish to use for payment.

**Book a Flight**

Round Trip  One Way

From: Eindhoven

To: Budapest-Terminal 1

Depart: 14 Aug 2008

Return: 03 Sep 2008

Passengers: 1 Adult Over 14 years  
0 Child 2 to 14 years

**going out** 

Eindhoven » Budapest-Terminal 1

date	fareclass	flight	departs	arrives	price excluding tax	taxes and charges
Fri 15 Aug 08	Web	W6 228	13:25	15:20	<input type="radio"/> Adult 94.99 EUR	26.00 EUR
Sun 17 Aug 08	Web	W6 228	13:25	15:20	<input type="radio"/> Adult 73.99 EUR	26.00 EUR

Next Week >  Previous Week <

***You have to find a hotel, so you look for...***

# ... a really cheap accommodation, or ...

Hostels - online booking at Youth Hostels and Backpackers Hostels - instant online reservations and reviews with HostelTraveler.co... X

File Edit View History Bookmarks Tools Help

[HostelTraveler.com](#) http://www.hosteltraveler.com/index.php

Welcome member travelers! [Sign in](#)

**HostelTraveler.com**

Hostels Reviews Best deals Top Cities

everything for travelers:

Search Hotels/Cities

Check Reservations

Free Membership

Hostel Traveler ... instant online booking for youth hostels, budget hotels and hostelling the globe

InstantWorldBooking

hostels Lodging Sign in lodging

**Find Hostels and Lodging at your destinations.**

Look for  for instant online booking.

► How it Works

► Step 1-Find Lodging

► Step 2-Make Reservations

secure online hostel booking at worldwide youth hostels, backpackers, and budget hotels.

Home 

Start Over

Members

Lodging Operators

About this Site

Selection Status

3 accommodations have been found matching your criteria.

Select your accommodations and click [Make Reservations](#) for rates, availability, and reservations.

Click on names to see photos, reviews, and more information.

**Tip:** Click  for rates and instant secure confirmations.

**Sort by:**  [Price \(Lo-Hi\)](#)  [Price \(Hi-Lo\)](#)  [Traveler Rating](#)  [Hostel Class](#)  
 [Hostel Name](#)

[View Advanced Display Options](#)

<b>Balaton, Hungary</b>		<a href="#">Make Reservations</a>
<input type="checkbox"/>	<b>Unity Hostel Balaton</b> Rakoczi Ut 268 Hostel 8 Units <a href="#">Write a Review</a>	  From <b>€12</b> 2hours from Budapest,we are located right behind a free beach access to the lake, right opposite a large shopping and dinning court, only few minutes from the best clubs.lots of freebies!!!

*... or a really luxurious one, or ...*

Search Results - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://www.hilton.com/en/hi/hotels/search/newresults.jhtml?searchType=city&stat

**Hilton**  
Travel should take you places®

Sign in

Username or HHonors #: Password or PIN:  
    
   Remember Me  
[View Text Only](#)

A prou  
TheHil  
be ho

Customer Support  
1-800-HILTONS

Specials & Packages Reservations Meetings Social Gatherings Hilton HHonors Things to Do eBrochures My Favo

Search Results

Hilton Hotels



Print Help

Change Your Search

Location

City: Budapest

Search Within: 40 mi  km

State/Province: State / Province

Country: Hungary

Brand

Hilton Hotels  All Hilton Family Hotels

Dates

Search Results

The following locations matched your request.  
**Rates displayed may be non-refundable.**  
 Please review rate rules and restrictions prior to booking your stay.

Sort by: Brand Show: All Hotels Go > View Hotels on a Map >

Hilton Hotels

**Hilton Budapest WestEnd**

Vaci ut 1-3  
Budapest, Hungary, 1062  
36-1-288-5500

Compare Hotel  Available

# ... an intermediate one ...

Szállás Utazás Magyarországon - Wellness-Szállás, Nyaralás, Utazás, Programok - Magyarországon - Mozilla Firefox

File Edit View History Bookmarks Tools Help

[http://www.travelsinhungary.hu/view\\_kat.php?catid=1&megye=budapest](http://www.travelsinhungary.hu/view_kat.php?catid=1&megye=budapest)

MAGYAR ENGLISH DEUTSCH Keresés Az összes kategóriában OK

**TRAVELS HUNGARY**  
www.travelsinhungary.hu

**TÖRZSVENDÉG KEDVEZMÉNYEK**

**REGISZTRÁCIÓ**

**ELÉRHETŐSÉG**

**VENDÉGKÖNYV**

Szállás (Wellness, Aktív pihenés, Gyógyturizmus)

Étterem

Programok, Látnivalók

Szórakozás (Élményfürdők, Kalandparkok)

Rendezvényszervezés (konferenciák-tréningek)

AJÁNLATOK

**Szálláshely \ Budapest**

Válasszon megyét:  
Válasszon!

Válasszon várost:  
Válasszon!

Ajánlatok száma a kategóriában: 1.





**Airport Hotel Budapest\*\*\*\***



**SZÁLLÁSHELY**  
**Budapest(Vecsés)**  
Ferihegyi repülőtéri szálloda és konferencia központ  
részletesen...

*oops, that is no good, the page is in  
Hungarian that almost nobody  
understands, but...*

*... this one could work*

Bestwestern.com, the World's Largest Hotel Chain - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://book.bestwestern.com/bestwestern/selectHotel.do?iata=00158210

Welcome to bestwestern.com®

Customer Service ↗ Rewards Program ↗ Gift Card ↗ Groups & Meetings ↗

**My Profile** Email or Member#:  SIGN-IN  
Forgot Password  
Create Password  
Enroll Now  
Password:

HOME FIND A HOTEL CHECK RESERVATIONS TRIP PLANNER PROMOTIONS & PROGRAMS PACKAGES

RESERVATIONS HOTEL LOCATIONS BEST WESTERN PREMIER HOTELS NEW HOTELS QUESTIONS & ANSWERS

→ Hotel Search Results ← Select Occupancy Select Room Review & Reserve Confirmation

Find a Hotel - Select Hotel Page: 1

Modify Your Search:

City: Budapest

Select State or Province: All

Select Country: Hungary

Check-In: 08/15/2008

Check-Out: 09/03/2008

Features & Amenities: High Speed Internet Complimentary Breakfast Breakfast Available

5 Hotels Found within 50 Mi / 80 Km of the Budapest Area

Show: All Hotels By: Distance

Display Currency In:

Show Hotels on a Map

**Best Western Hotel Hungaria**

Stay at this 4-star Budapest hotel offering guests deluxe accommodations near some of Budapest's popular attractions and business locations. Visitors... More >

  
Photo Gallery

Rákóczi Ut 90,  
Budapest, H-1074, Hungary  
Distance from City Center: 0.86m / 1.38km

Hotel.net

Hotel not available on selected dates.  
Check Alternate Dates

*Of course, you could decide to trust a specialized site...*

*... like this one, or...*

Create your package from Amsterdam to Budapest (and vicinity) - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://www.expedia.com/pub/agent.dll?qscr=cmsh&subm=1&CMBTX\_0\_rgnm=Bud

Welcome - Already a member? Log In | Sign Up | My Itineraries | My Account | Customer Support

Search Expedia

Expedia

Home Flights Hotels Cars Vacation Packages Cruises Activities DEALS & OFFERS Maps Business Travel Rewards

Start search over

Change your search

Departing: 8/18/2008

Returning: 9/3/2008

Star Rating: Show all

Lodging Type: Show All

Name contains:

Go

Maps: Area map view Hotel map view

Hotel amenities: Narrow your search

Show hotels in this area: Budapest (and vicinity) (all areas) Go

Not what you're looking for? Choose a different destination

Create your package from Amsterdam to Budapest (and vicinity)

View packages: 1 - 25 26 - 50 51 - 75 76 - 97 Previous | Next

Sort by:  Expedia Picks  Price  Hotel Name  City  Star Rating  Traveler Opinion

 Corinthia Grand Hotel Royal Avg/person: \$2889

Impressive landmark building with imposing Neo-classical façade and soaring glass atrium, set on Pest's busy Erzsébet Avenue, and housing shops, a spa, ... More lodging info

Hotel promotion - Stay 3 Nights and Save 20% on Your Stay!

★★★★★ Budapest, PEST

Executive Double-Executive lounge usage Check in: 8/18/08

Includes: Free Wireless Internet, Spa Credit, Breakfast Buffet Check out: 9/03/08

Traveler Opinion 4.7 out of 5 15+ reviews

Amsterdam (AMS) to Budapest (BUD) Depart: 8/18/08 6:00 PM - 8:00 PM Malev Hungarian Airlines

1-800-555-XXXX Call now for the same great deals plus expert advice

*... or this one*

Budapest Hotels: Read Budapest Hotel Reviews and Compare Prices - TripAdvisor - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://www.tripadvisor.com/Hotels-g274887-Budapest-Hotels.html

My TripAdvisor Register Now!

22,582,888 Travelers from 190 Countries Planned Trips Here This Week!

Home Destinations Fun & Games Just For Members

Home → Europe → Hungary → Budapest

## Budapest Hotels

ADD TO MY MAP WRITE A REVIEW EMAIL

Hotels (272) B&Bs / Inns (24) Specialty Lodging (83)

Find Hotels Travelers Trust

Check-in: Any date mm/dd/yyyy Check-out: Any date mm/dd/yyyy My dates are flexible

Price: Any Price U.S. Dollars Hotel class: Any Class Adults: 2

Recommended For: All

Find Hotels

Best deals: Budapest hotels

- Budapest: Boek en bespaar tot 75%. Booking.com Geen reserveringskosten!
- Great Budapest Hotels www.Venere.com/Budapest\_Hotels See maps & pics, read book online. Relax and enjoy your stay!
- Budapest easyHotel Deals www.easyHotel.com City Centre from just €15 per night from founder of easyJet
- Cheap Hotels Budapest www.otel.com/BudapestHotels Fantastic rates on Budapest! Huge Savings, Instant Confirmation!

View all deals for Budapest

Recommended Hotels (1-20 of 272)

Sort by: Popularity next >

Free Budapest Guide Get the best picks for where to eat and play.

Download pdf

*You may want to know something about  
Budapest; look for some photographs...*

*... on flickr ...*

Flickr: Budapest - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://www.flickr.com/groups/budapest/

Signed in as Ivan Herman

flickr LOVES YOU

Home You Organize Contacts Groups Explore

Search this group's pool Search

Budapest

Group Pool | Discussion | 1,418 Members | Map | Invite Friends

Share This

Group Pool ( 19,017 items | Add photos or video )

 NEW From apuc

 NEW From André Fromont

 NEW From carlogambino

 NEW From carlogambino

 NEW From Crashbandi

 NEW From Crashbandi

 NEW From Crashbandi

 NEW From Crashbandi

 NEW From Crashbandi

 NEW From Crashbandi

 NEW From Crashbandi

 NEW From Crashbandi

» More

Discussion ( 33 posts | Post a new topic )

# ... on Google ...

budapest - Google Image Search - Mozilla Firefox

File Edit View History Bookmarks Tools Help

<http://images.google.nl/images?hl=en&q=budapest&btnG=Search+Images&gbv=1>

Web Images Maps News Video Gmail more ▾ Sign in

**Google™**  [Search Images](#) [Search the Web](#) [Advanced Image Search Preferences](#)

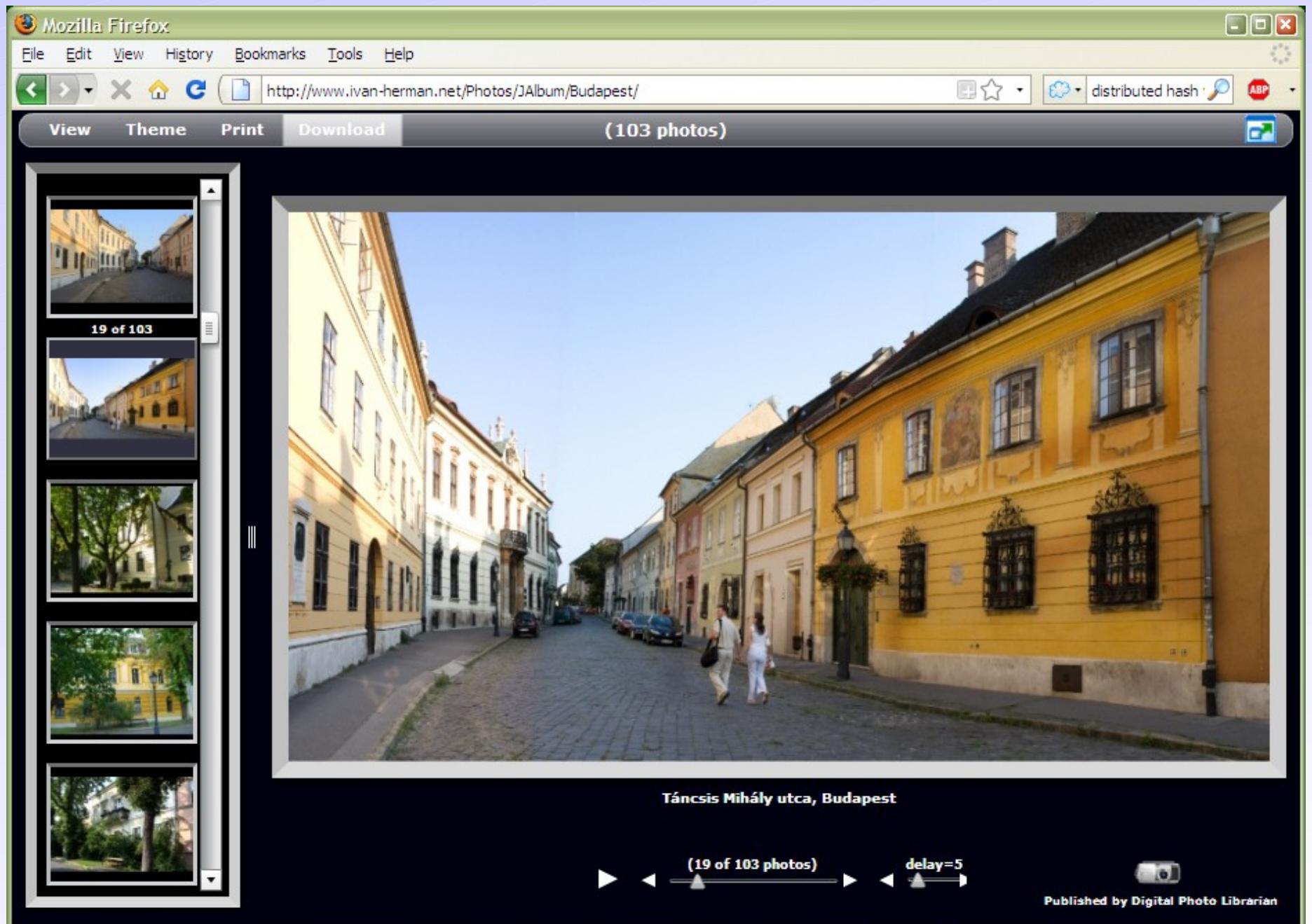
Moderate SafeSearch is on

Images Showing: All image sizes

Results 1 - 18 of about 19,900,000 for budapest [definition]. (0.25 seconds)

					
E-mail this photo E-mail. Budapest 550 x 412 - 40k - jpg <a href="http://www.tripadvisor.com">www.tripadvisor.com</a>	Budapest, Hungary 450 x 338 - 43k - jpg <a href="http://www.transitionsabroad.com">www.transitionsabroad.com</a>	Budapest looks its most beautiful at ... 1024 x 768 - 161k - jpg <a href="http://web.kvif.bgf.hu">web.kvif.bgf.hu</a>	Beautiful-Budapest 430 x 320 - 34k - jpg <a href="http://www.budapesthotels.com">www.budapesthotels.com</a> [ More from <a href="http://www.budapesthotels.com">www.budapesthotels.com</a> ]	Hotel Victoria Budapest 575 x 473 - 92k - jpg <a href="http://www.victoria.hu">www.victoria.hu</a>	Fly to Budapest and experience one ... 909 x 682 - 347k - jpg <a href="http://www.sterling.dk">www.sterling.dk</a>
					
Danube Bridge Elisabeth in Budapest ... 1024 x 768 - 194k - jpg <a href="http://budapest5.freeblog.hu">budapest5.freeblog.hu</a>	Budapest had 2421831 inhabitants in ... 422 x 426 - 29k - jpg <a href="http://www.squidoo.com">www.squidoo.com</a>	budapest night 575 x 352 - 206k - jpg <a href="http://www.wayfaring.info">www.wayfaring.info</a>	Budapest - Things to Do with a Day ... 400 x 300 - 32k - jpg <a href="http://cruises.about.com">cruises.about.com</a>	Jewish Cultural Heritage in Budapest 452 x 360 - 87k <a href="http://www.budapesthotels.com">www.budapesthotels.com</a>	Hungary, Budapest, Parliament 640 x 480 - 199k - jpg <a href="http://www.hungary.travelphotoguide.co">www.hungary.travelphotoguide.co</a>

*... or you can look at mine* 😊



# ... or a (social) travel site

Budapest Travel Guide | Budapest Tourism - RealTravel - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://realtravel.com/budapest-hungary-travel-guide-d13081-1.html travel tom ABP

realtravel™ Real People. Real Advice. Real Experiences.™ Free Travel Blog Free Trip Planner Sign In

Travel Guides Hotels Attractions Things To Do Editor's Picks Deals Search

You are here: Destinations > Europe > Hungary > Budapest > Travel Guide

BUDAPEST TRAVEL GUIDE AND TOURISM

Introduction Guides Trips Photos Flights Hotels Restaurants Attractions Deals

 Budapest Travel Guide

This capital city - made up of two parts, Buda and Pest - sits on one of the most beautiful areas of the Danube River and it shows. Often dubbed the "Paris of Eastern Europe", it is a combination of Old World grandeur and a thriving cultural scene. Budapest has a vibrancy and vitality that never slows and the numerous sights can occupy travelers for weeks. With so much history and culture to explore, no traveler leaves unsatisfied.

photo by Taistea

more Budapest photos ▶ Destinations near Budapest ▶

Travel Guide Information From Our Partners

# What happened here?

- You had to consult a large number of sites, all different in style, purpose, possibly language...
- You had to mentally *integrate* all those information to achieve your goals
- We all know that, sometimes, this is a long and tedious process!

- All those pages are only tips of respective icebergs:
  - the real *data* is hidden somewhere in databases, XML files, Excel sheets, ...
  - you have only access to what the Web page designers allow you to see

- Specialized sites (Expedia, TripAdvisor) do a bit more:
  - they gather and combine data from other sources (usually with the approval of the data owners)
  - but they still control how you see those sources
- But sometimes you want to personalize: access the original data and combine it yourself!

# Here is another example...

**CoCoDat - Collation of Cortical Data - Mozilla Firefox**

File Edit View History Bookmarks Tools Help

CoCoMac DATABASES ORT EXAMPLES

**Cell Centered Database - Mozilla Firefox**

File Edit View History Bookmarks Tools Help

http://ccdb.ucsd.edu/sand/main?event=gallery&action=show&dpl=y

**Cell Centered Database™**  
National Center for Microscopy and Imaging Research **Gallery**

Data | Search | Gallery | Dictionary | Publications | MyCCDB | Data Download | Contact us | Help

2D image Reconstruction Segmentation Animation

**CoCoDat: Collation of Cortical [sic] microcircuitry] Data**

CoCoDat is a microcircuitry database that published experimental reports. The data and cellular compartment), as well as the

- Morphology
- Firing properties
- Ionic currents
- Ionic conductances
- Synaptic currents
- Connectivity

The database is available for download under data tables but also a Search Board with p manual or automatic relaxation of the sea

- Brain region
- Layer
- Neuron type

<http://www.cocomac.org/cocodat/catalyzer/index.html>

**NeuronDB - Thalamic relay neuron - Overview (A) () - Mozilla Firefox**

File Edit View History Bookmarks Tools Help

http://senselab.med.yale.edu/ senselab

**NeuronDB**

Back Thalamic relay neuron

Mode: **Overview** Data/Search plus Connectivity plus Classical References/Notes Models

Region: Distal equivalent dendrite Middle equivalent dendrite Proximal equivalent dendrite Soma Axon hillock Axon fiber Axon terminal All Compartments

Properties: Receptors Channels Transmitters **All Properties**

Interoperation: Gene and Chromosome Experimental Data (neurodatabase.org) Microscopy Data (CCDB)

Neuron type: principal Organism: Vertebrates

1. Equivalent dendrite  
2. Distal equivalent dendrite  
3. Middle equivalent dendrite  
4. Proximal equivalent dendrite  
5. Soma

Show other Show other Show other Show other Show other Show other

Done Z PIP logged out

*Another example: social sites. I have a list of “friends” by...*

# ... Dopplr,

DOPPLR: Ivan Herman's fellow travellers - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://www.dopplr.com/traveller/IvanHerman/fellows

Find and Invite | Your connections | Your trips | Your account | Sign out  
Visit our blog for updates

DOPPLR FOR IVAN

Where Next? { Gent, Boston, Vienna...  
+ Add a trip

Type the name of a city or a traveller

Find people and places

Welcome, Ivan

In the last 2 weeks,  
one of your fellow travellers added a trip that coincides with you.

You are at home in Amsterdam.

Find out more in your journal...

You can invite people to Dopplr to see your trips, find them on other networks you use  
or look for travellers you already know to encourage more coincidences.

You have a public profile. Edit it?

You can now create a public profile to display to the whole internet if you want, not just Dopplr users — and take  
any of the information to embed on your own website. Give it a try!

Your trips Fellow travellers Your journal Your carbon List or Map

Peter Brown is in Montréal until August 16th. Boston soon. Montréal later.

Daniel Appelquist is in Washington until August 23rd. Aspen soon. Washington later.

Eva Méndez is in Maraña until August 17th. Santo Domingo later.

Danny Weitzner is in Bergen until September 5th. Los Angeles later.

Charlton Barreto is in Vienna until August 18th. Sacramento soon. Vienna later.

# ... Twine,

My Connections | Twine - Mozilla Firefox

File Edit View History Bookmarks Tools Help

BETA INVITE-ONLY twine

http://www.twine.com/user/ivanherman/connections

Home My Items My Twines My Connections Explore Start a Twine

Manherman Logout Account

Search Twine

## My Connections

Search within My Connections...

View All Sort by Most Recent

Refine your view by... You can also filter by selecting from the following categories.

► related twines

  
Dean Allemand  
at home  
Oakland, CA  
3 Twines | 4 Items  
[Send Message](#) | [Disconnect](#)

  
David Provost  
Breathing  
Boston  
20 Twines | 15 Items  
[Send Message](#) | [Disconnect](#)

  
Attila Gardos  
Hungary, Budapest  
9 Twines | 8 Items  
[Send Message](#) | [Disconnect](#)

  
jim  
got my san fra  
83 Twi

  
Steve  
Seattle  
38 Twi

  
James  
living in  
95008  
122 Twi

  
Chris  
All rea  
Mill Val  
73 Twi

# ... LinkedIn,

LinkedIn: My Contacts: Connections - Mozilla Firefox

File Edit View History Bookmarks Tools Help

[http://www.linkedin.com/connections?trk=hb\\_side\\_cnts](http://www.linkedin.com/connections?trk=hb_side_cnts)

Account & Settings | Help | Sign Out Language ▾

Advanced Search People Search

LinkedIn

People | Jobs | Answers | Companies

Home Groups Profile Contacts Inbox (15)

Add Connections

Ivan Herman  
Semantic Web Activity Lead, World Wide Web Consortium  
What are you working on?  
Your profile is 80% complete [Edit]

Aasman, Jans CEO at Franz Inc  
View & edit details » 252

Abramatic, Jean-François Chief Product Officer at ILOG  
View & edit details » 163

Adida, Ben Software Security Researcher and Entrepreneur  
View & edit details » 148

Allemang, Dean Chief Scientist at TopQuadrant Inc.  
View & edit details » 179

Allison, Kevin San Francisco Correspondent at The Financial Times  
View & edit details » 152

Alonso, Jose Manuel eGovernment Lead at W3C  
View & edit details » 106

Connections Imported Contacts Network Statistics Add Connections | F

Showing 311 of 311 connections. 15 outstanding sent Invitations

A B C D E F G H I J K L M N O P Q R S T

# ... and, of course, Facebook

Facebook | All Friends - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://www.new.facebook.com/friends/#/friends/?flid=0&view=everyone& easy jet

Welcome to the new Facebook [Send feedback](#) | [Back to the old Facebook](#)

facebook Ivan Herman Friends Applications Inbox Home Settings Logout Search

All Friends

**Friend Lists**

- All Friends
- CWI
- Family
- IW3C2
- W3C Team

[Make a New List](#)

**Find Friends**

- [Find people you know who are already on Facebook](#)
- [Invite your friends to join Facebook](#)

Showing: [Status Updates](#) [Recently Updated](#) [Phonebook](#) **Everyone** [Search Friends](#)

Show: [Choose an option...](#)

You have 137 friends. [1](#) [2](#) [3](#) [Next](#)

	<a href="#">Shadi Abou-Zahra</a>	
	<a href="#">Ross Ackland</a>	
	<a href="#">Ben Adida</a>	
	<a href="#">Anupriya Ankolekar</a>	
	<a href="#">Daniel Appelquist</a> is hanging out in Aspen. 15 hours ago	
	<a href="#">Lora Aroyo</a>	

Advertise

Discover Ali world n°1 o marketplace

Do You T Photos? Make your c book, from

More Ads

- I had to type in and connect with friends again and again for each site independently 😞
- This is even worse then before: I feed the icebergs, but I still do not have an easy access to data...

# What would we like to have?

- Use the data on the Web the same way as we do with documents:
  - be able to link to data (independently of their presentation)
  - use that data the way I want (present it, mine it, etc)
  - agents, programs, scripts, etc, should be able to *interpret* part of that data

## *Put it another way...*

- We would like to *extend* the current Web to a “Web of data”:
  - allow for applications to exploit the data directly

***But wait! Isn't what mashup sites are already doing?***

# A “mashup” example:

**TripIt | Organize your travel - Mozilla Firefox**

File Edit View History Bookmarks Tools Help

Summary View Expanded View

Budapest, Hungary (Edit)  
Avg: Sunny / Hi 28°C / Lo 14°C

View Change Log

**Trip Details**

Thursday, July 24, 2008

**Flight from Amsterdam (AMS) to Zurich (ZRH)**

**FLIGHT**  
14:55 CEST

**Swiss International Airlines 729**

Depart: Amsterdam (AMS), 14:55 CEST  
nonstop 1h 25min  
aircraft Airbus A320-100/200  
Arrive: Zurich (ZRH), 16:20 CEST 374 miles

Passengers: Ivan Herman, Eva Boka ep Herman

Booking Information: Booked on 18/4/2008 <http://www.swiss.com>

**Flight from Zurich (ZRH) to Budapest (BUD)**

**FLIGHT**  
17:20 CEST

**Swiss International Airlines 2258**

Depart: Zurich (ZRH), 17:20 CEST  
nonstop 1h 35min  
aircraft Fokker 100  
Arrive: Budapest (BUD), 18:55 CEST 500 miles

Passenger(s): Ivan Herman, Eva Boka ep Herman.

**Map of Budapest, Hungary**

Budapest, Hungary

Map Satellite Hybrid

**MAPS**  
19:05 CEST

- In some ways, yes, and that shows the huge power of what such Web of data provides
- But mashup sites are forced to do very ad-hoc jobs
  - various data sources expose their data via Web Services
  - each with a different API, a different logic, different structure
  - these sites are forced to reinvent the wheel many times because there is no standard way of doing things 😞

# *Put it another way (again)...*

- We would like to extend the current Web to a standard way for a “Web of data”

# *But what does this mean?*

- What makes the current (document) Web work?
  - people create different documents
  - they give an address to it (ie, a URI) and make it accessible to others on the Web

# *Steven's site on Amsterdam (done for some visiting friends)*

The Internet Guide to Amsterdam - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://homepages.cwi.nl/~steven/amsterdam.html

The Internet Guide to Amsterdam



**Contents**

- [Introduction](#)
- [Time](#)
- [Weather](#)
- [Language](#)
- [Money](#)
- [Tipping](#)
- [Electricity](#)
- [Safety and Health](#)
- [Hotels](#)
- [Eating and Drinking](#)
- [Transport](#)
- [Shopping](#)
- [News](#)
- [Communications](#)
- [Places to See](#)
- [What's On](#)
- [The Amsterdam Year](#)
- [Maps](#)
- [Books](#)
- [Other Resources](#)

**Designed to be printed out and taken with you.**

Written by Steven Pemberton, CWI, Amsterdam, and Astrid Kerssens, Amsterdam.

Linked to by more than 450 other sites; more than **3,500,000** grunted readers!

*The top Amsterdam travel guide according to Google. If you know how Google works, you know that that says something about this site!*

See also [London](#)

**A Review of This Guide**

The Internet Guide To Amsterdam:

WebUser Gold Award

Rating: **★★★★★**

Reviewed By: Jane Hoskyn

Sometimes, it's the simple things in life that make your heart skip a beat.

This is especially true when you're abroad and you need a guide that loads really fast on your mobile phone or PDA (check), doesn't involve clicking from page to page to find what you want (check), isn't written in fluffy tourist-board-speak (check), has all the links you need and none of the ones you don't (check), and is put together by someone who loves your destination and knows it better than the insides of his own eyelids (check). Step forward Steven Pemberton, creator of this exemplary guide to having a damn good time in the Dam.

**Introduction**

Amsterdam is an unusual city in that it has all the advantages of a big city – culture, history, food, entertainment, good

## *Then some magic happens...*

- Others discover the site and they link to it
- The more they link to it, the more important and well known the page becomes
  - remember, this is what, eg, Google exploits!
- This is the “Network effect”: some pages become important, and others begin to rely on it even if the author did not expect it...

# This could be expected...

The screenshot shows a screenshot of an Opera browser window displaying the [WWW9 Organizers](http://www9.org/w9-organizers.html) website. The page title is "WWW9 Organizers" and the subtitle is "May 15-19, 2000, Amsterdam". A sidebar on the left contains links: "WWW9 HOME", "PROGRAM INFO", "SPONSORING", "EXHIBITING", "VOLUNTEERS", "ORGANIZERS", "PAST CONFERENCES", and "AMSTERDAM". The "AMSTERDAM" link is circled in red. The main content area is titled "WWW9 Conference Committee". It lists "Conference Co-Chairs": Ivan Herman, CWI, The Netherlands and Albert Vezza, CNRI, USA. It lists "Program Committee Chair": Dick Butlerman, Oratrix, The Netherlands. It lists "Program Committee Vice Chairs": Ann Bassetti, Boeing, USA; Stephan Fischer, Technical University, Darmstadt, Germany; and Lynda Hardman, CWI, The Netherlands. The browser's toolbar and status bar are visible at the bottom.

*but this one, from the other side of the Globe,<sup>42</sup>  
was not...*

Netherlands - Spring Break Information - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://msass.case.edu/international/SPNetherlandFAQ.html

CASE.EDU: HOME | DIRECTORIES | SEARCH

**CASE WESTERN RESERVE UNIVERSITY**  
MANDEL SCHOOL OF APPLIED SOCIAL SCIENCES

INTERNATIONAL COURSES/PROGRAMS

MSASS | About | Admissions | Programs | Students | Faculty | Research | Library | Calendar | Departments

**International Courses**

General Information  
Descriptions by Country  
Passport Application  
Dutch Visa Application  
Previous Trips  
Testimonials

**Netherlands - Spring Break Information**

Please also see the following links for information on the trip to the Netherlands

- General FAQ
- Applying for a Dutch Visa
- A Students Photo Journal of the Netherlands

**What Forms do I need for this Program?**

All of the forms you will need are linked from the Forms page.

More information? Check out these links:

<b>Amsterdam Links</b>
Spring Break Trip participants to Amsterdam may familiarize themselves with Amsterdam by visiting the following web sites. On website address <a href="#">nl</a> refers to the <a href="#">Netherlands</a> .
<i>Information on travel outside the United States , including instructions on getting a passport: <a href="http://travel.state.gov">http://travel.state.gov</a></i>
<i>A must see website <a href="http://homepages.cwi.nl/~steven/amsterdam.html">http://homepages.cwi.nl/~steven/amsterdam.html</a></i>
<i>United States Consulate, Amsterdam: <a href="http://www.usemb.nl">http://www.usemb.nl</a></i>
<i>Virtual Tour of Amsterdam: <a href="http://www.channels.nl">http://www.channels.nl</a></i>

# What would that mean for a Web of Data?

- Lessons learned: we should be able to:
  - “publish” the data to make it known on the Web
    - standard ways should be used instead of ad-hoc approaches
    - the analogous approach to documents: *give URI-s to the data*
  - make it possible to “link” to that URI from *other* sources of data (not only Web pages)
    - ie, applications should not be forced to make targeted developments to access the data
    - generic, standard approaches should suffice
  - and let the network effect work its way...

# *But it is a little bit more complicated* 😕

- On the traditional Web, humans are implicitly taken into account
- A Web link has a “context” that a person may use

# Eg: address field on my page:

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: [\(sha1sum: 5ac8032d5f6012ea1775ea2f63e1676bafd5e80b\)](mailto:ivan@w3.org)

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 GnuPG 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:



# ... leading to this page

Centrum Wiskunde & Informatica | CWI - Mozilla Firefox

File Edit View History Bookmarks Tools Help

<http://www.cwi.nl/>

Centrum Wiskunde & Informatica

home | contact | nl | internet | Search...

CWI

about cwi events library news research scientists

**Centrum Wiskunde & Informatica**

Centrum Wiskunde & Informatica (CWI) performs fundamental scientific research in mathematics and computer science. CWI transfers the acquired knowledge to society and industry. The institute's strategy for the period up to 2012 is to concentrate research on four broad, societally relevant themes:

**Earth & life sciences** **The data explosion** **Societal logistics** **Software as service**

---

**News**

13-10-08  
PhD defence Jarek Byrka  


22-09-08  
**Best Paper Award for SMIL State research**  
At the ACM DocEngineering Symposium in São Paulo, Brazil, from 16 till 19 September, CWI researchers Jack Jansen and Dick Bulterman received the Best Paper Award.  
[read more](#)

Agenda

18-10-08  
**Science Day at the Science Park Amsterdam**  
At Science Day the Science Park Amsterdam will be open for This year the theme of Science Day is 'Crack the code'. CWI workshops en demonstrations within this theme for every age  
[read more](#)

21-10-08  
**MAS Seminar, speaker Svetlana Dubinkina**  
Two speaker session  
Tea starting at 10.00  
Room: M279  
Speakers:  
1. Svetlana Dubinkina, CWI MAS 1, tba  
2. James Glazier, University of Indiana Bloomington, tba  
[read more](#)

28-10-08  
**MAS Seminar, speaker Peter Sonneveld**

- A human understands that this is my institution's home page
- He/she knows what it means (realizes that it is a research institute in Amsterdam)
- On a Web of Data, something is missing; machines can't make sense of the link alone

- New lesson learned:

- extra information (“label”) must be added to a link: “this links to my institution, which is a research institute”
- this information should be machine readable
- this is a *characterization* (or “classification”) of *both* the link *and* its target
- in some cases, the classification should allow for some limited “reasoning”

# *Let us put it together*

- What we need for a Web of Data:
  - use URI-s to publish data, not only full documents
  - allow the data to link to other data
  - characterize/classify the data and the links (the “terms”) to convey some extra meaning
  - and use standards for all these!

# ***So What is the Semantic Web?***

*It is a collection of standard technologies  
to realize a Web of Data*

- It is that simple...
- Of course, the devil is in the details
  - a common model has to be provided for machines to describe, query, etc, the data and their connections
  - the “classification” of the terms can become very complex for specific knowledge areas: this is where ontologies, thesauri, etc, enter the game...



## *In what follows...*

- We will use a simplistic example to introduce the main technical concepts
- The details will be for later during the course

# *The rough structure of data integration*

1. Map the various data onto an abstract data representation
  - make the data independent of its internal representation...
2. Merge the resulting representations
3. Start making queries on the whole!
  - queries that could not have been done on the individual data sets

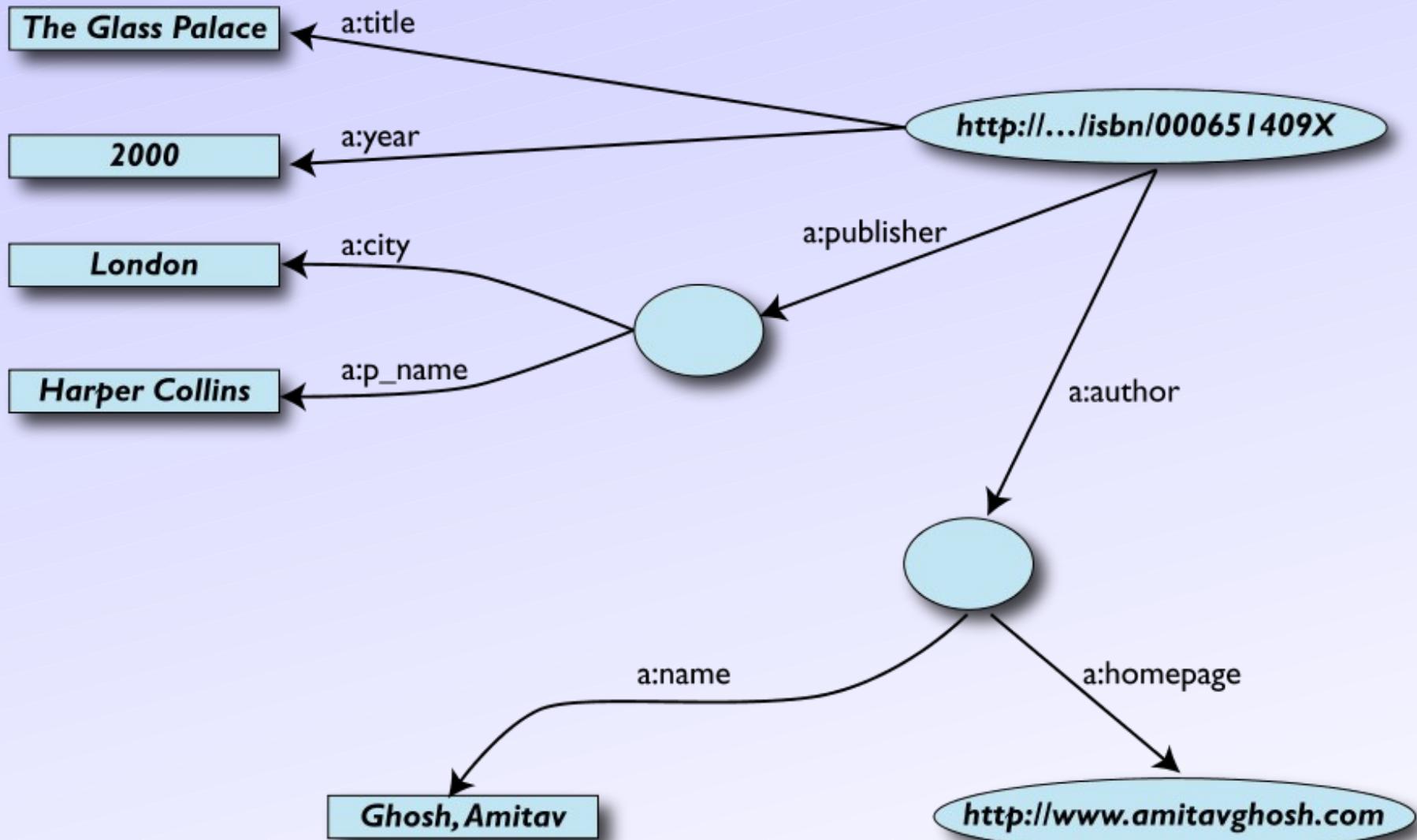
# A *simplified* bookstore data (dataset “A”)

<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

# 1<sup>st</sup>: export your data as a set of relations



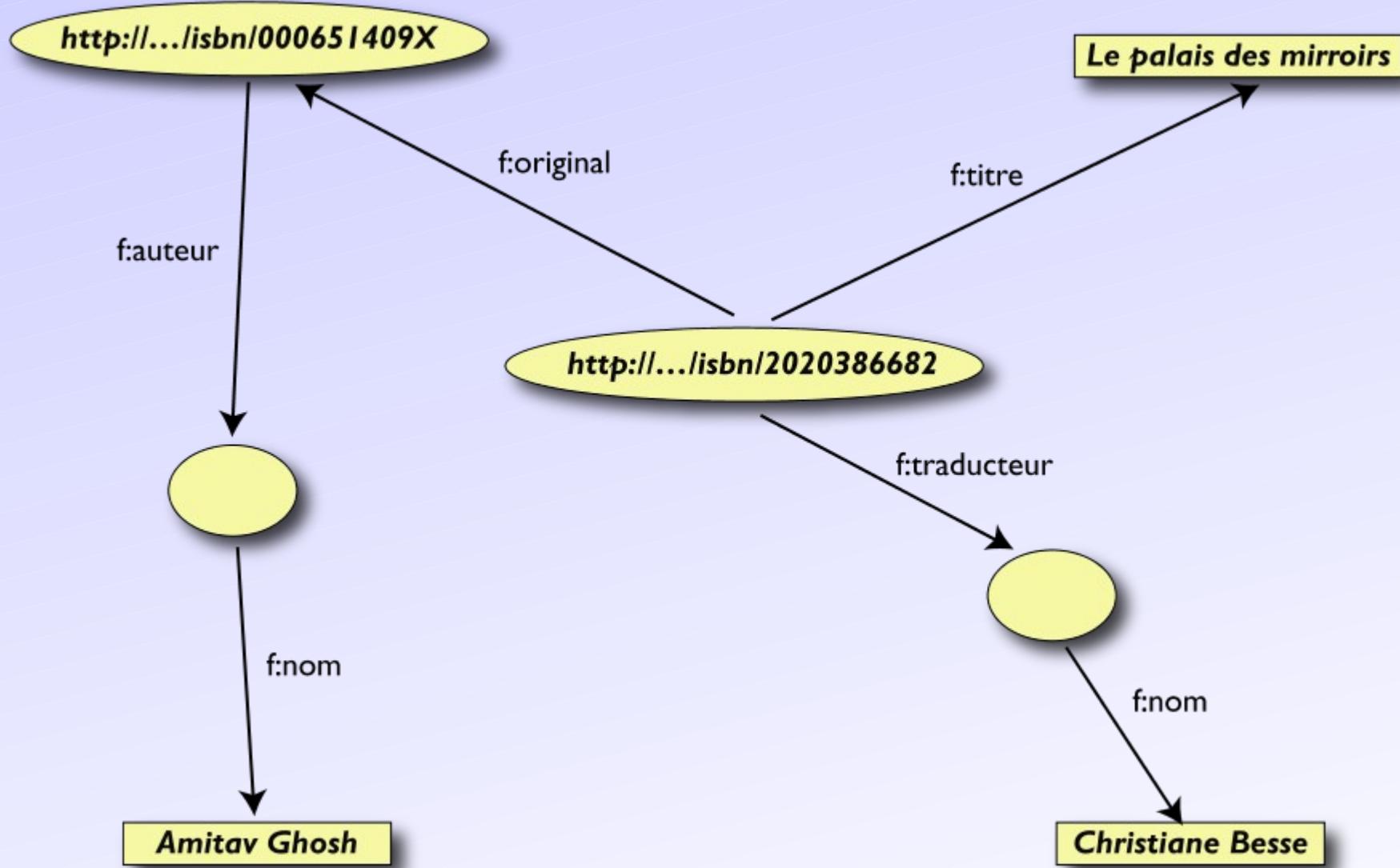
# *Some notes on the exporting the data*

- Relations form a graph
  - the nodes refer to the “real” data or contain some literal
  - how the graph is represented in machine is immaterial for now
- Data export does not necessarily mean physical conversion of the data
  - relations can be generated on-the-fly at query time
    - via SQL “bridges”
    - scraping HTML pages
    - extracting data from Excel sheets
    - etc.
- One can export part of the data

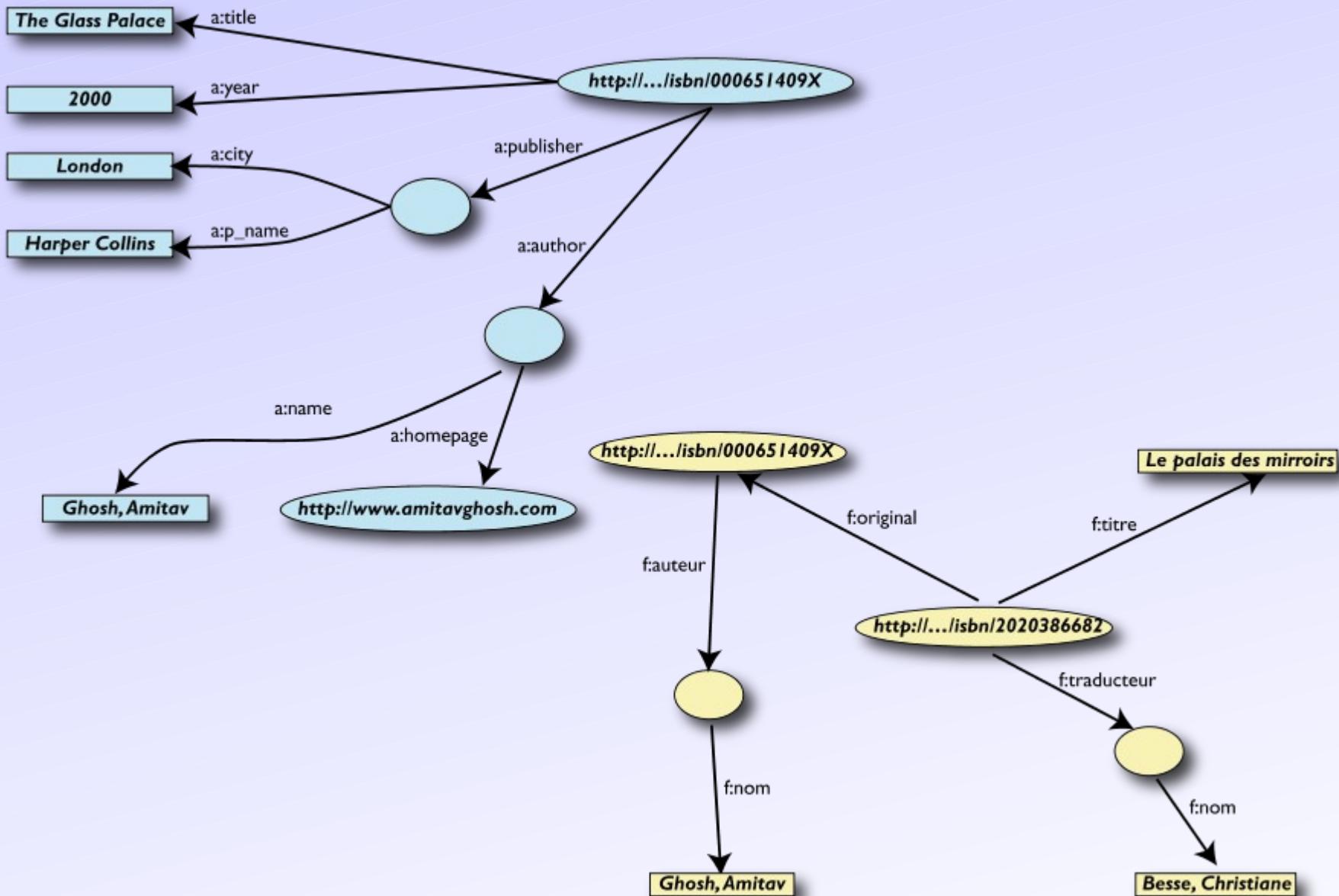
# Another bookstore data (dataset “F”)

	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		

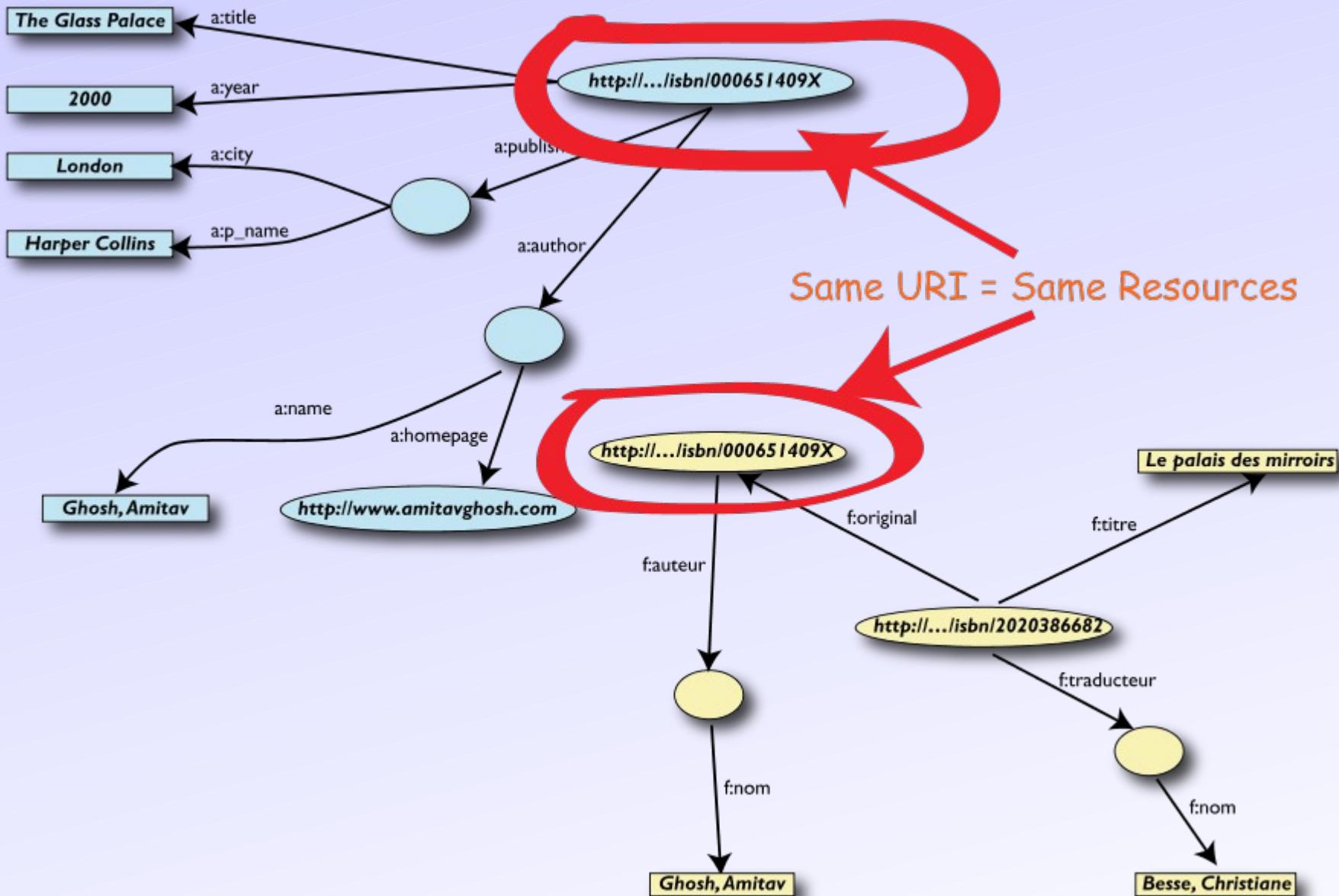
## 2<sup>nd</sup>: export your second set of data



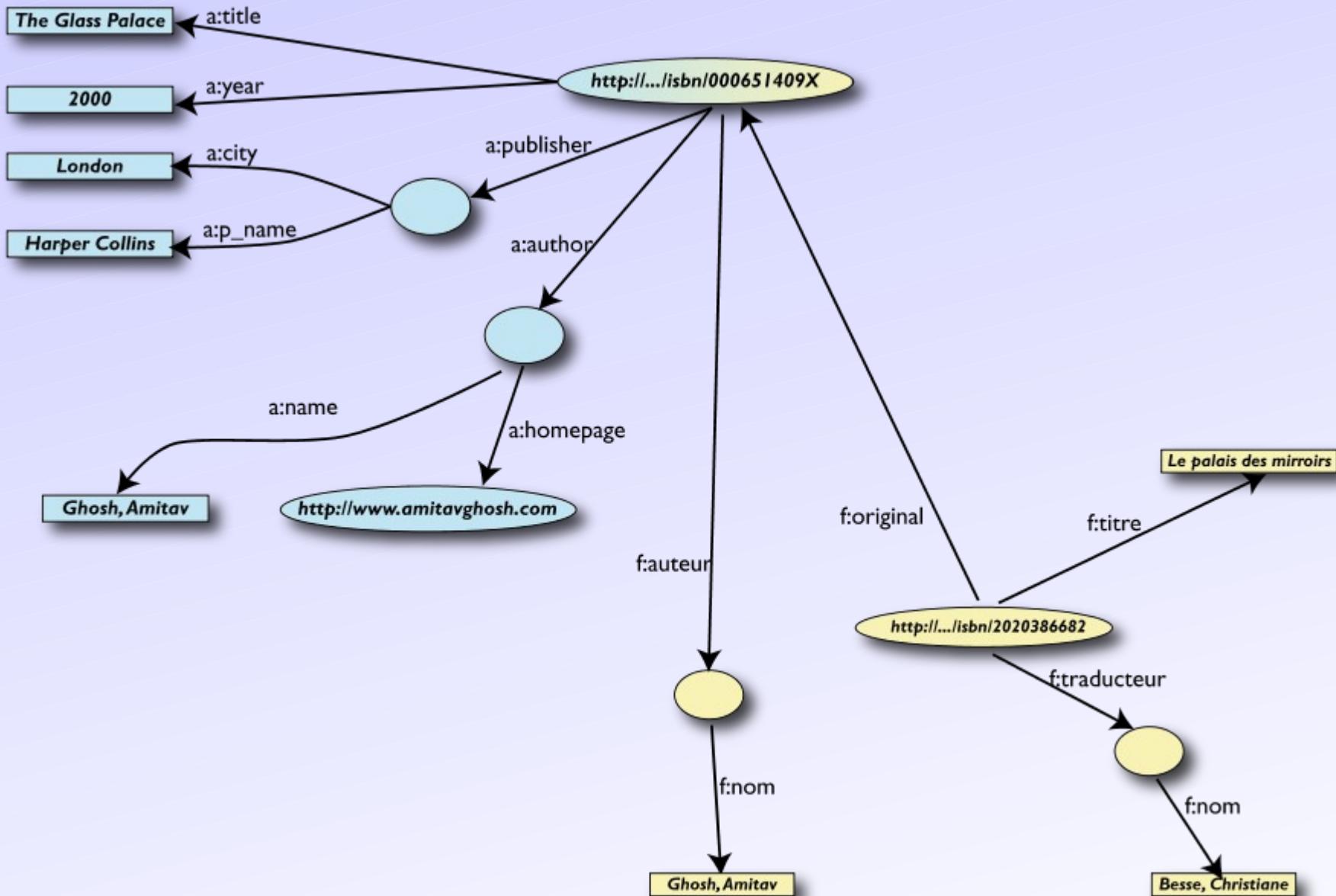
# 3<sup>rd</sup>: start merging your data



# 3<sup>rd</sup>: start merging your data (cont.)

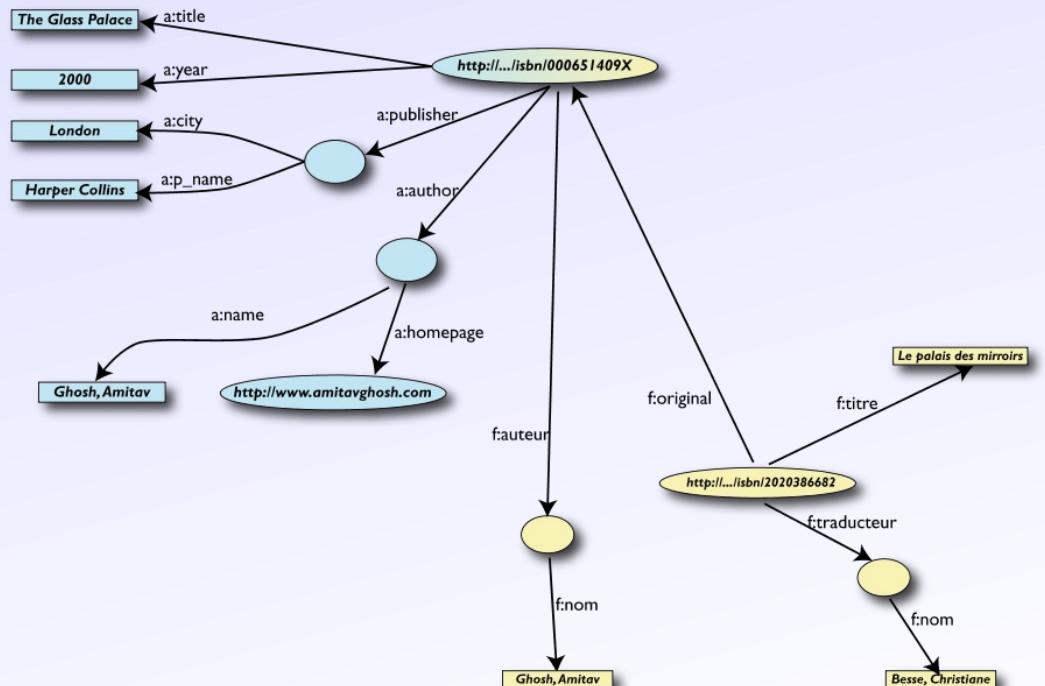


# 3<sup>rd</sup>: merge identical resources



# Start making queries...

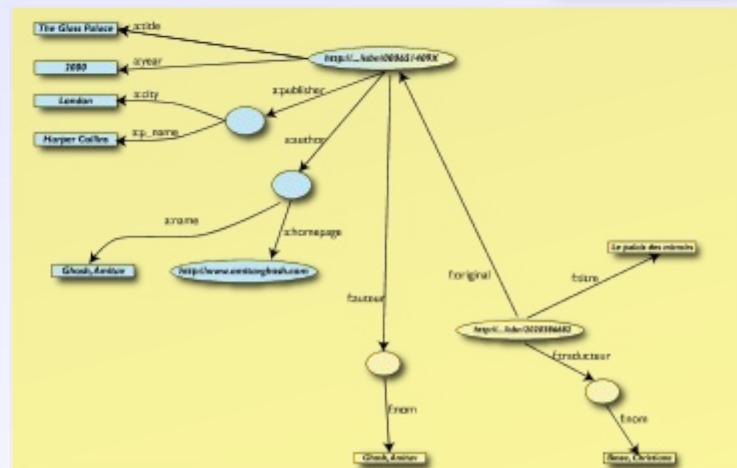
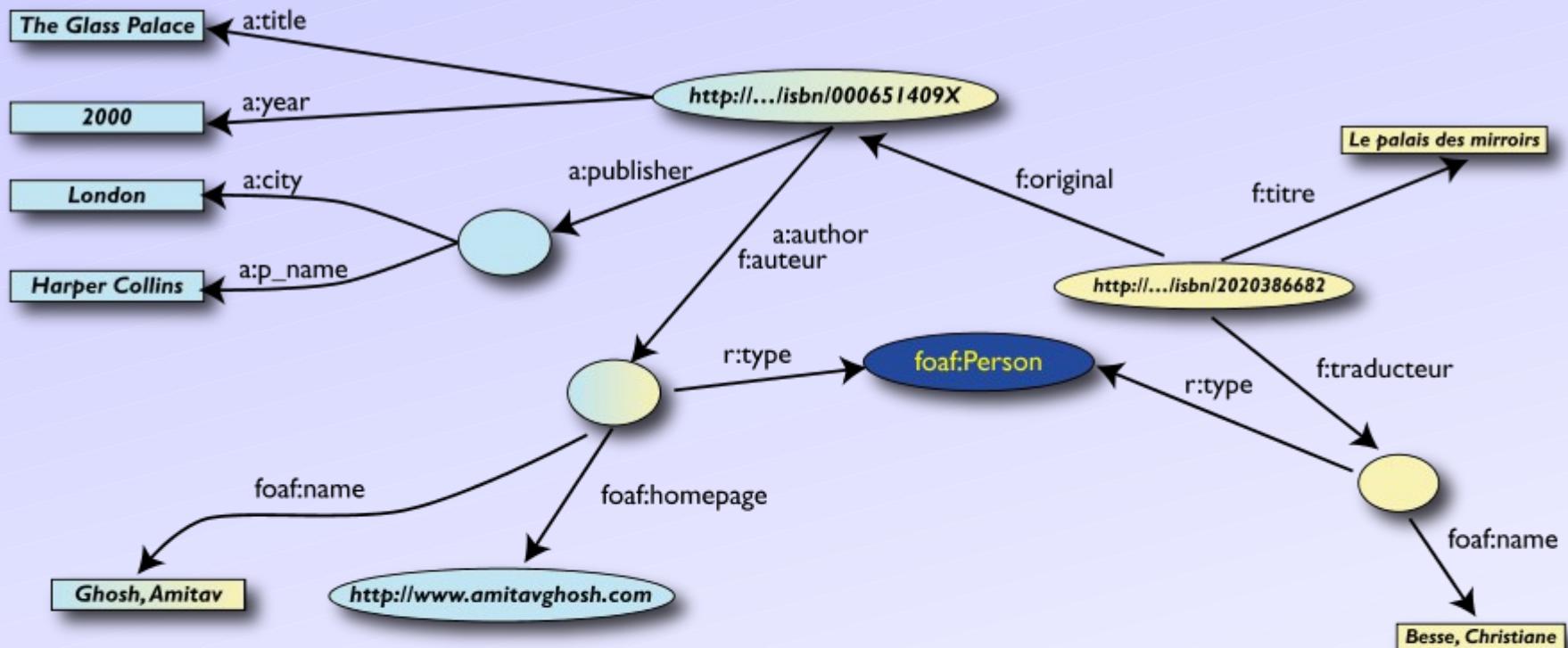
- User of data “F” can now ask queries like:
  - “give me the title of the original”
    - well, ... « donne-moi le titre de l’original »
- This information is not in the dataset “F”...
- ...but can be retrieved by merging with dataset “A”!



## *However, more can be achieved...*

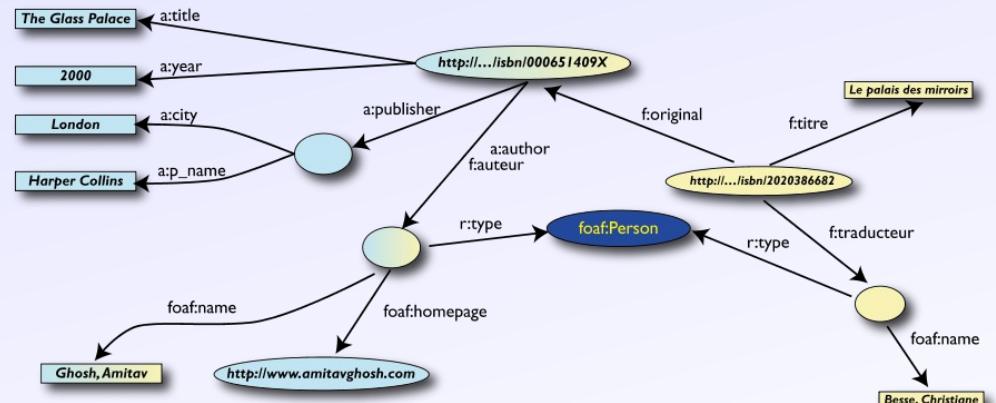
- We “feel” that **a:author** and **f:auteur** should be the same
- But an automatic merge does not know that!
- Let us add some extra information to the merged data:
  - **a:author** same as **f:auteur**
  - both identify a “Person”
  - a term that a community may have already defined:
    - a “Person” is uniquely identified by his/her name and, say, homepage
    - it can be used as a “category” for certain type of resources

# 3<sup>rd</sup> revisited: use the extra knowledge



# *Start making richer queries!*

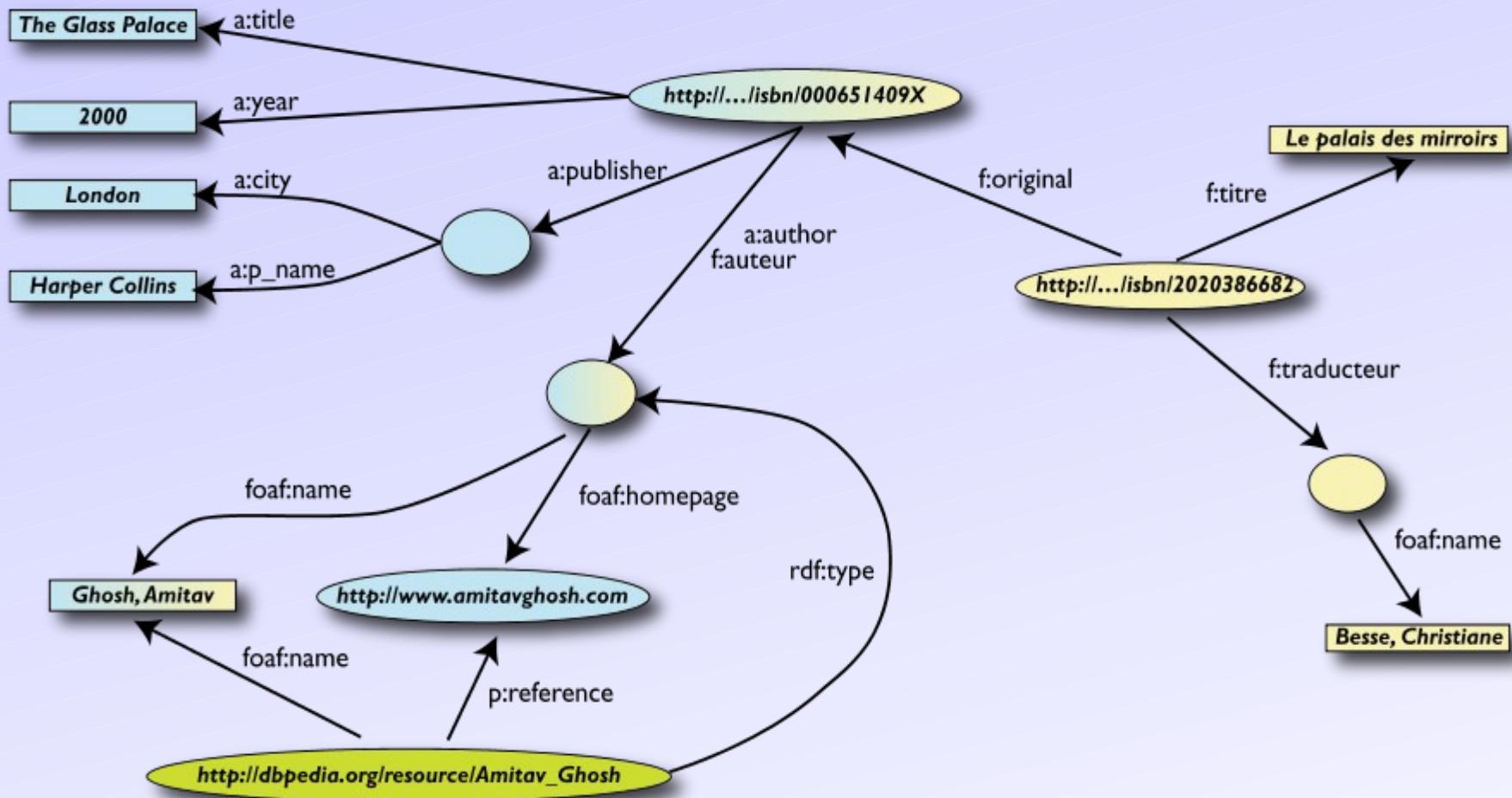
- User of dataset “F” can now query:
  - “donnes-moi la page d'accueil de l'auteur de l'originale”
    - well... “give me the home page of the original’s ‘auteur’”
- The information is not in datasets “F” or “A”...
- ...but was made available by:
  - merging datasets “A” and datasets “F”
  - adding three simple extra statements as an extra “glue”



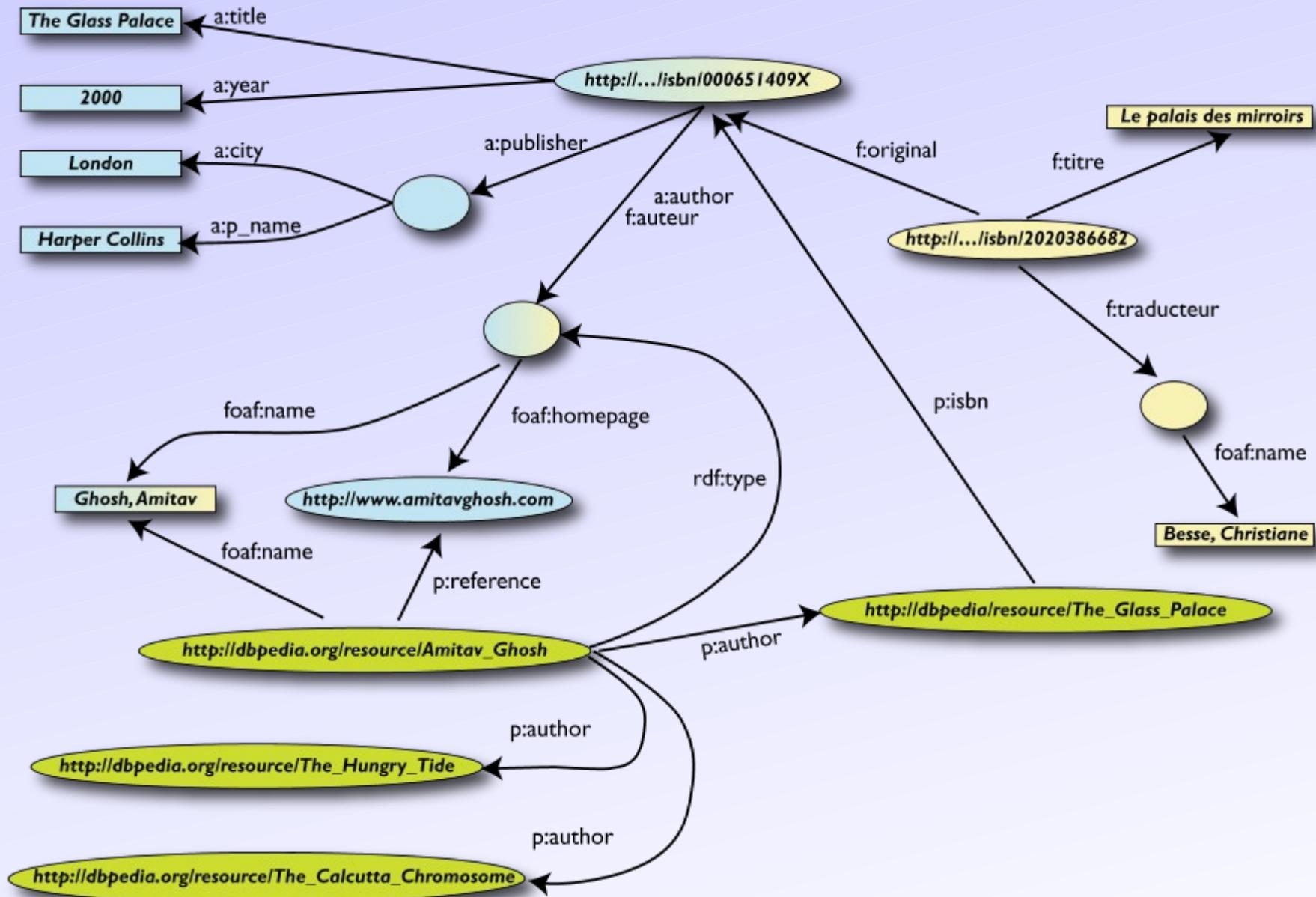
# **Combine with different datasets**

- Using, e.g., the “Person”, the dataset can be combined with other sources
- For example, data in Wikipedia can be extracted using dedicated tools
  - e.g., the “[dbpedia](#)” project can extract the “infobox” information from Wikipedia already...

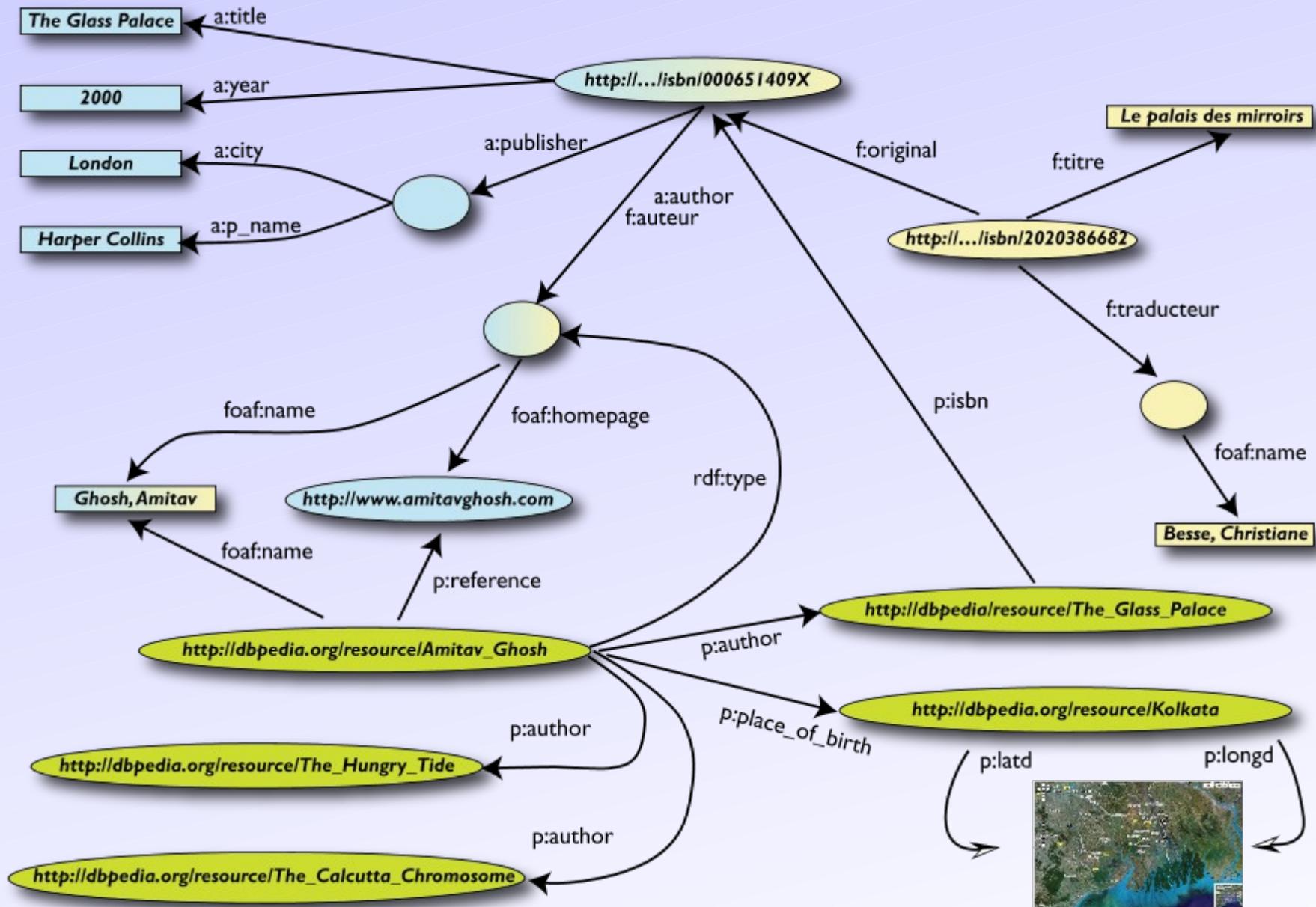
# Merge with Wikipedia data



# Merge with Wikipedia data



# Merge with Wikipedia data



# *Is that surprising?*

- It may look like it but, in fact, it should not be...
- What happened via automatic means is done every day by Web users!
- The difference: a bit of extra rigour so that machines could do this, too

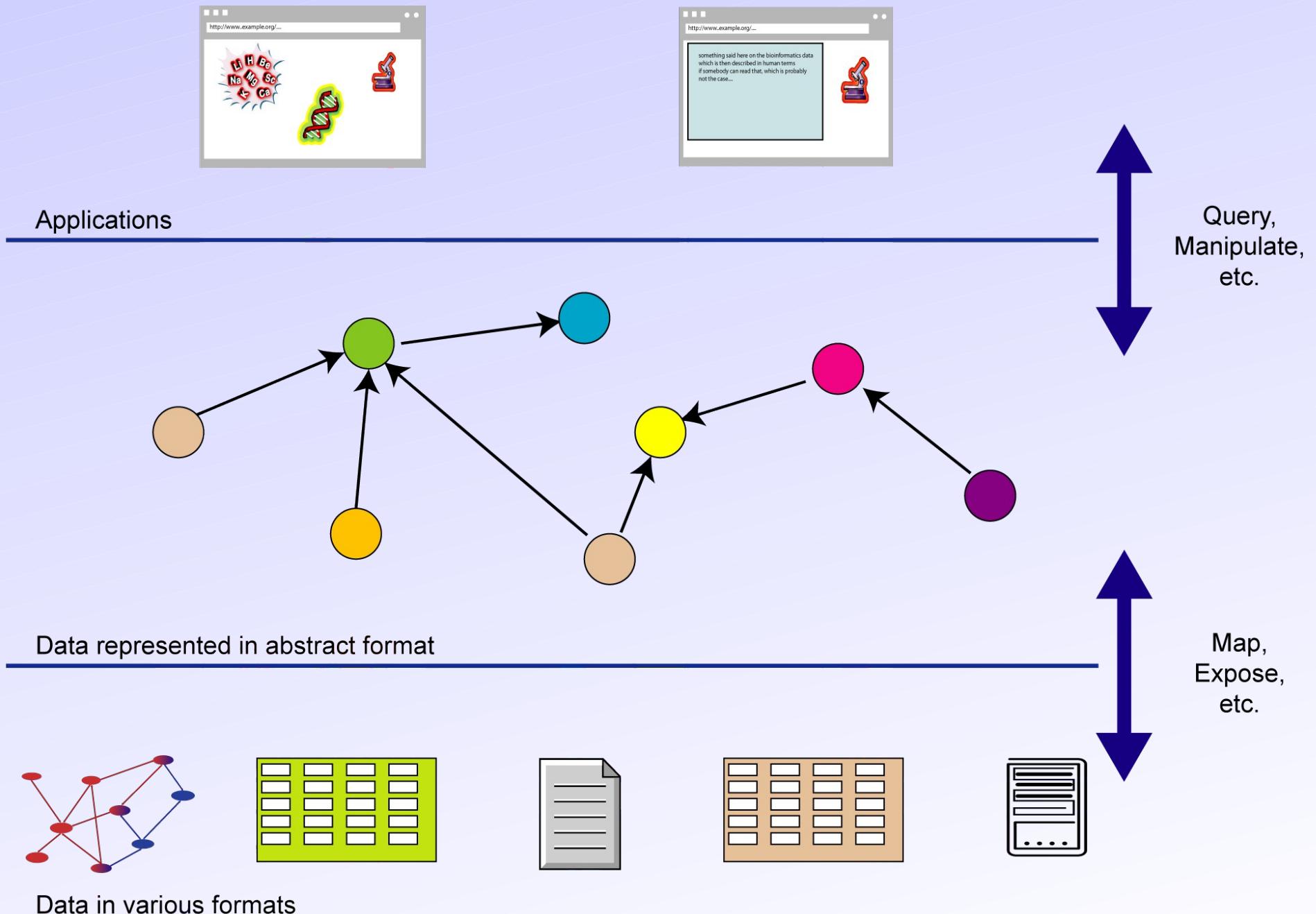
# What did we do?

- We combined different datasets that
  - are somewhere on the web
  - are of different formats (mysql, excel sheet, XHTML, etc)
  - have different names for relations
- We could combine the data because some URI-s were identical (the ISBN-s in this case)
- We could add some simple additional information (the “glue”), possibly using common terminologies that a community has produced
- As a result, new relations could be found and retrieved

# ***It could become even more powerful***

- We could add extra knowledge to the merged datasets
  - e.g., a full classification of various types of library data
  - geographical information
  - etc.
- This is where ontologies, extra rules, etc, come in
  - ontologies/rule sets can be relatively simple and small, or huge, or anything in between...
- Even more powerful queries can be asked as a result

# What did we do? (cont)



# *The Basis: RDF*

# RDF triples

- Let us begin to formalize what we did!
  - we “connected” the data...
  - but a simple connection is not enough... data should be named somehow
  - hence the RDF Triples: a labelled connection between two resources

## RDF triples (cont.)

- An RDF Triple ( $s, p, o$ ) is such that:
  - “ $s$ ”, “ $p$ ” are URI-s, ie, resources on the Web; “ $o$ ” is a URI or a literal
    - “ $s$ ”, “ $p$ ”, and “ $o$ ” stand for “subject”, “property”, and “object”
  - here is the complete triple:

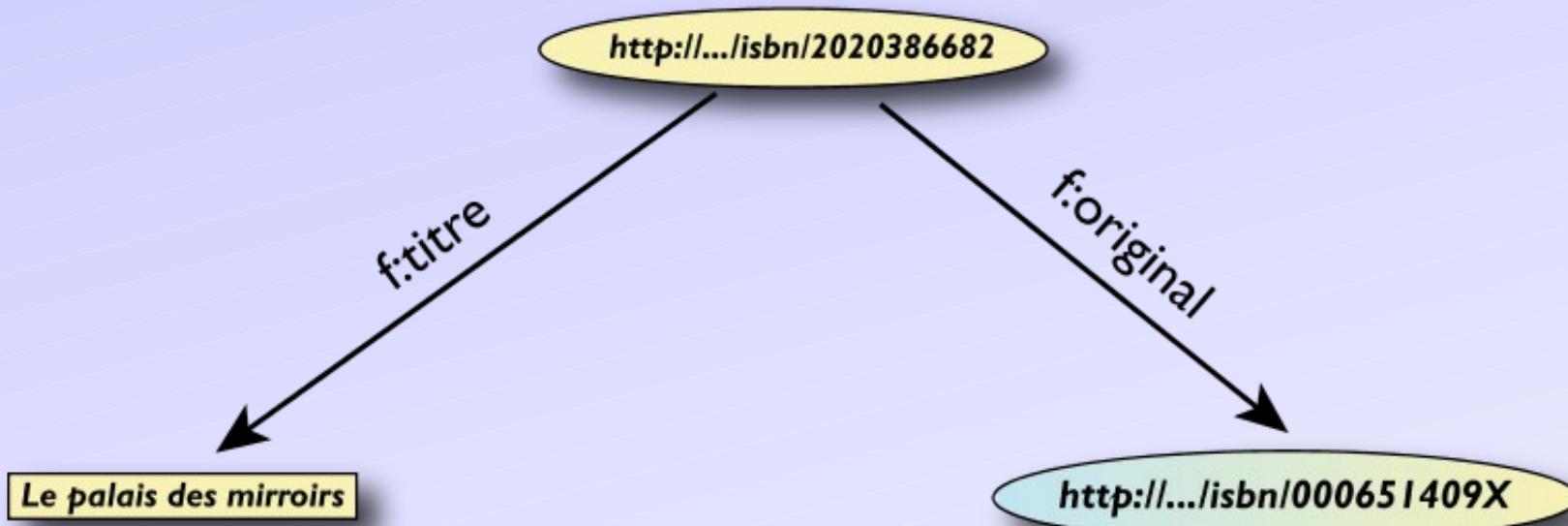
```
(<http://...isbn...6682>, <http://.../original>, <http://...isbn...409x>)
```

- RDF is a general model for such triples (with machine readable formats like RDF/XML, Turtle, N3, RXR, ...)

# RDF triples (cont.)

- Resources can use *any* URI, e.g.:
  - `http://www.example.org/file.xml#element(home)`
  - `http://www.example.org/file.html#home`
  - `http://www.example.org/file2.xml#xpath1(//q[@a=b])`
- URI-s can also denote non Web entities:
  - `http://www.ivan-herman.net/me` is me
    - not my home page, not my publication list, but me
- RDF triples form a directed, labelled graph

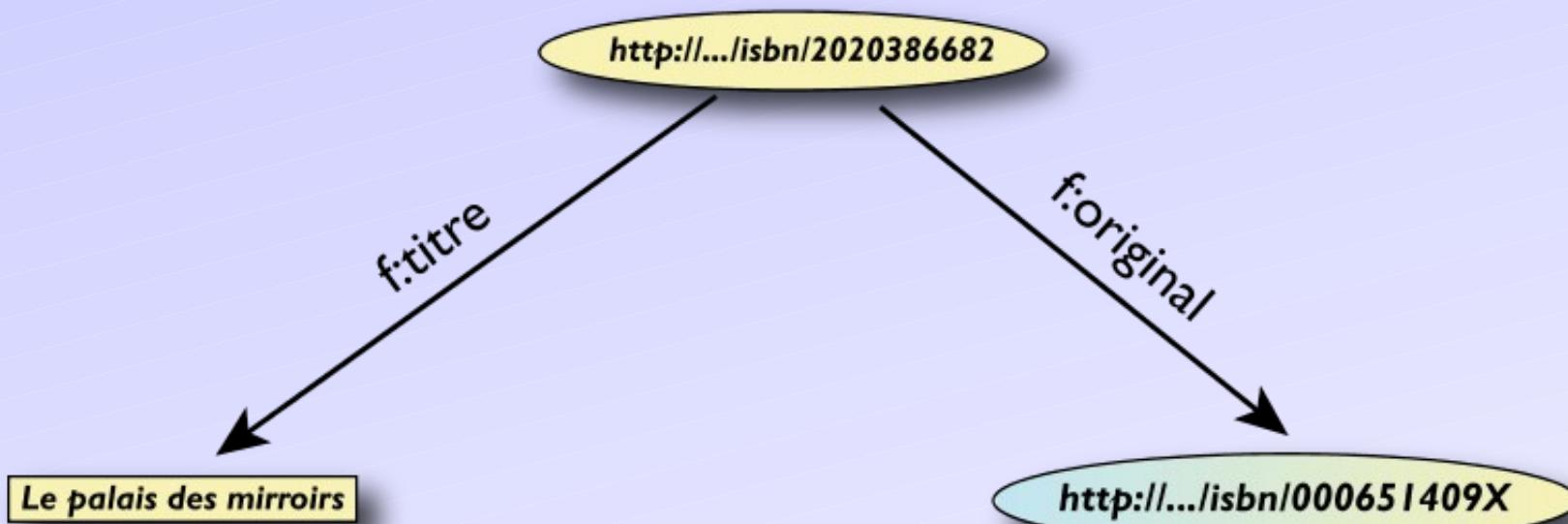
# A simple RDF example (in RDF/XML)



```
<rdf:Description rdf:about="http://.../isbn/2020386682">
  <f:titre xml:lang="fr">Le palais des mirroirs</f:titre>
  <f:original rdf:resource="http://.../isbn/000651409X"/>
</rdf:Description>
```

(Note: namespaces are used to simplify the URI-s)

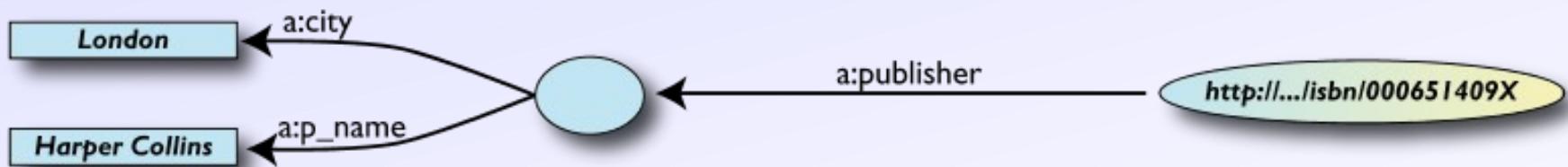
# A simple RDF example (in Turtle)



```
<http://.../isbn/2020386682>
  f:titre "Le palais des mirroirs"@fr ;
  f:original <http://.../isbn/000651409X> .
```

# “Internal” nodes

- Consider the following statement:
  - “the publisher is a «thing» that has a name and an address”
- Until now, nodes were identified with a URI. But...
- ...what is the URI of «thing»?



# *Internal identifier (“blank nodes”)*

```
<rdf:Description rdf:about="http://.../isbn/000651409X">
  <a:publisher rdf:nodeID="A234"/>
</rdf:Description>
<rdf:Description rdf:nodeID="A234">
  <a:p_name>HarpersCollins</a:p_name>
  <a:city>HarpersCollins</a:city>
</rdf:Description>
```

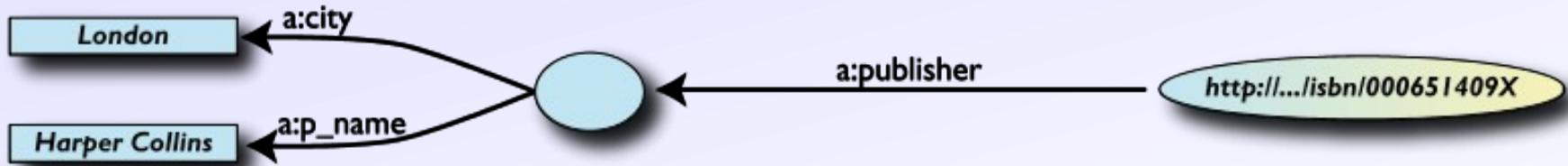
```
<http://.../isbn/2020386682> a:publisher _:A234 .
_:A234 a:p_name "HarpersCollins".
```

- Syntax is serialization dependent
- A234 is invisible from outside (it is not a “real” URI!); it is an internal identifier for a resource

# Blank nodes: the system can also do it

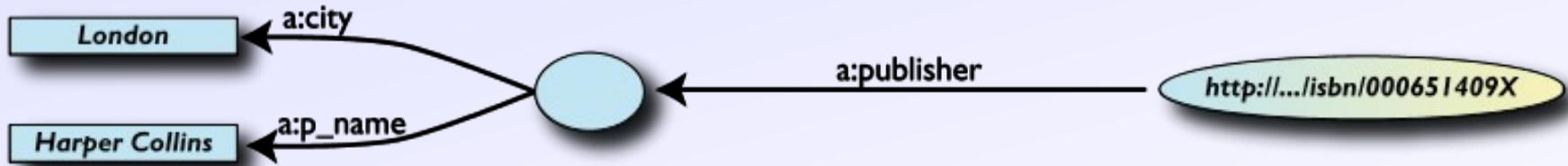
- Let the system create a “nodeID” internally (you do not really care about the name...)

```
<rdf:Description rdf:about="http://.../isbn/000651409X">
  <a:published>
    <rdf:Description>
      <a:p_name>HarpersCollins</a:p_name>
      ...
    </rdf:Description>
  </a:published>
</rdf:Description>
```



# Same in Turtle

```
<http://.../isbn/000651409X> a:publisher [  
    a:p_name "HarpersCollins";  
    ...  
].
```



## Blank nodes: some more remarks

- Blank nodes require attention when merging
  - blanks nodes with identical nodeID-s in different graphs are different
  - implementations must be careful...
- Many applications prefer not to use blank nodes and define new URI-s “on-the-fly”

# *RDF in programming practice*

- For example, using Java+Jena (HP's Bristol Lab):
  - a “Model” object is created
  - the RDF file is parsed and results stored in the Model
  - the Model offers methods to retrieve:
    - triples
    - (property,object) pairs for a specific subject
    - (subject,property) pairs for specific object
    - etc.
  - the rest is conventional programming...
- Similar tools exist in Python, PHP, etc.

# Jena example

```
// create a model
Model model=new ModelMem();
Resource subject=model.createResource("URI_of_Subject")
// 'in' refers to the input file
model.read(new InputStreamReader(in));
StmtIterator iter=model.listStatements(subject,null,null);
while(iter.hasNext()) {
    st = iter.next();
    p = st.getProperty();
    o = st.getObject();
    do_something(p,o);
}
```

## *Merge in practice*

- Environments merge graphs automatically
  - e.g., in Jena, the Model can load several files
  - the load merges the new statements automatically

# Example: integrate experimental data

- Goal: reuse of older experimental data
- Keep data in databases or XML, just export key “fact” as RDF
- Use a faceted browser to visualize and interact with the result

**Internal Compound Repurposing Example**

Welcome, Allergy & Respiratory Team Member

This tool allows you to identify opportunities for additional uses of compounds from other teams within your project. It combines internal data, public data and the results of data mining experiments to provide testable hypotheses.

Control Panel & Item Filtering					
Area	5: Approach	3: Term+Reason	1: Max_Stage_Reached	1: Literature Links	
29 Pain	<input checked="" type="checkbox"/> 7: Antibody	<input type="checkbox"/> 3: ACTIVE	<input type="checkbox"/> 51: Candidate	<input checked="" type="checkbox"/> 0 - 50	
16 Metabolic Disease	<input checked="" type="checkbox"/> 1: Recombinant	<input type="checkbox"/> 12: BIOMARKER	<input type="checkbox"/> 10: Discovery	<input type="checkbox"/>	
3: Cancer	<input type="checkbox"/> 18: SM_Agonist	<input checked="" type="checkbox"/> 51: EFFICACY	<input checked="" type="checkbox"/> 4: Exploratory	<input type="checkbox"/>	

**Internal Compound Repurposing Example**

Welcome, Allergy & Respiratory Team Member

This tool allows you to identify opportunities for additional uses of compounds from other teams within your project. It combines internal data, public data and the results of data mining experiments to provide testable hypotheses.

Control Panel & Item Filtering					
Area	5: Approach	3: Term+Reason	1: Max_Stage_Reached	1: Literature Links	
29 Pain	<input checked="" type="checkbox"/> 7: Antibody	<input type="checkbox"/> 3: ACTIVE	<input type="checkbox"/> 51: Candidate	<input checked="" type="checkbox"/> 0 - 50	
16 Metabolic Disease	<input checked="" type="checkbox"/> 1: Recombinant	<input type="checkbox"/> 12: BIOMARKER	<input type="checkbox"/> 10: Discovery	<input type="checkbox"/>	
3: Cancer	<input type="checkbox"/> 18: SM_Agonist	<input checked="" type="checkbox"/> 51: EFFICACY	<input checked="" type="checkbox"/> 4: Exploratory	<input type="checkbox"/>	
3: Sexual Health	<input checked="" type="checkbox"/> 12: SM_Antagonist	<input type="checkbox"/> 11: MARKET	<input type="checkbox"/> 19: HTS	<input type="checkbox"/>	
2: Infectives	<input checked="" type="checkbox"/> 21: SM_Inhibitor	<input type="checkbox"/> 11: REORG	<input type="checkbox"/> 11: Phase I	<input type="checkbox"/>	
1: Urogenitals	<input checked="" type="checkbox"/>	<input type="checkbox"/> 10: TOXIC	<input type="checkbox"/> 13: Phase III	<input type="checkbox"/>	
			<input type="checkbox"/> 41: Screening	<input type="checkbox"/>	

51 items filtered from 710 originally (Reset All Filters)

Area	Original + Indication	Target_Name	Approach	Start	Term+Reason	Max_Stage_Reached	Owner	OBIM Lit_All Lit_2007 Lit_Mech IMA GEO Pathway Compounds
Metabolic Disease	Diabetes	Liver glycogen phosphorylase	SM_Inhibitor	2007-Q2	EFFICACY	Candidate	P. Person	SW-030072
Sexual Health	Erectile Dysfunction	Integrin alpha-3 (Glycoprotein 33/MuA3) (CD49c)	SM_Antagonist	2006-Q3	EFFICACY	Candidate	P. Person	SW-029782
Sexual Health	Erectile Dysfunction	Leukotriene C4 synthase	SM_Agonist	2006-Q3	EFFICACY	Candidate	M. Manager	SW-029638
Sexual Health	Erectile Dysfunction	transcription elongation factor A (SII)-like 4	SM_Inhibitor	2005-Q2	EFFICACY	Candidate	P. Person	SW-029926
Infectives	HIV	Putative four-repeat ion channel (Ix)	SM_Inhibitor	2006-Q2	EFFICACY	Candidate	L. Leader	SW-029994
Infectives	HIV	Voltage-gated potassium channel protein KV.2 (Ix)	SM_Agonist	2007-Q1	EFFICACY	Candidate	A. Scientist	SW-029653
Urogenital Incontinence		Human RNA binding motif (RBM) gene, partial cdk.	SM_Agonist	2007-Q3	EFFICACY	Candidate	L. Leader	SW-029684
Pain	Migraine	Monocarboxylate transporter homologue294064CD1 (Ix)	SM_Inhibitor	2007-Q3	EFFICACY	Candidate	L. Leader	SW-030085

# *One level higher up*

*(RDFS, Datatypes)*

# *Need for RDF schemas*

- First step towards the “extra knowledge”:
  - define the terms we can use
  - what restrictions apply
  - what extra relationships are there?
- Officially: “RDF Vocabulary Description Language”
  - the term “Schema” is retained for historical reasons...

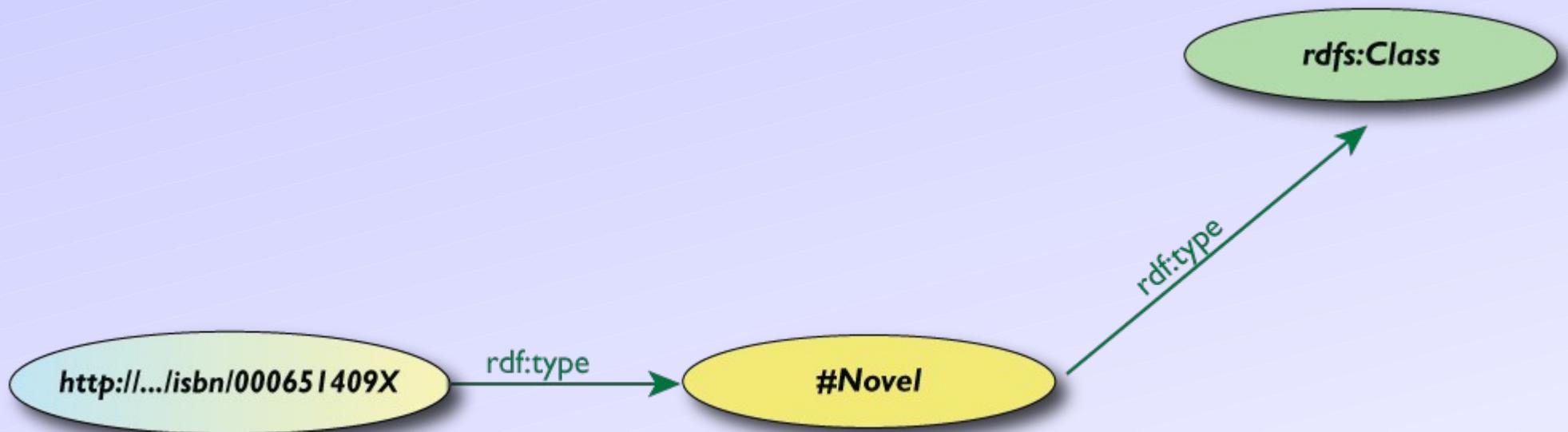
# *Classes, resources, ...*

- Think of well known traditional ontologies or taxonomies:
  - use the term “novel”
  - “every novel is a fiction”
  - “«The Glass Palace» is a novel”
  - etc.
- RDFS defines resources and classes:
  - everything in RDF is a “resource”
  - “classes” are also resources, but...
  - ...they are also a collection of possible resources (i.e., “individuals”)
    - “fiction”, “novel”, ...

# Classes, resources, ... (cont.)

- Relationships are defined among classes and resources:
  - “typing”: an individual belongs to a specific class
    - “«The Glass Palace» is a novel”
    - to be more precise: “«<http://.../000651409x>» is a novel”
  - “subclassing”: *all* instances of one are also the instances of the other (“every novel is a fiction”)
- RDFS formalizes these notions in RDF

# Classes, resources in RDF(S)



- RDFS defines the meaning of these terms
  - (these are all special URI-s, we just use the namespace abbreviation)

# *Schema example in RDF/XML*

- The schema part:

```
<rdf:Description rdf:id="Novel">
  <rdf:type
    rdf:resource="http://www.w3.org/2000/01/rdf-schema#Class"/>
</rdf:Description>
```

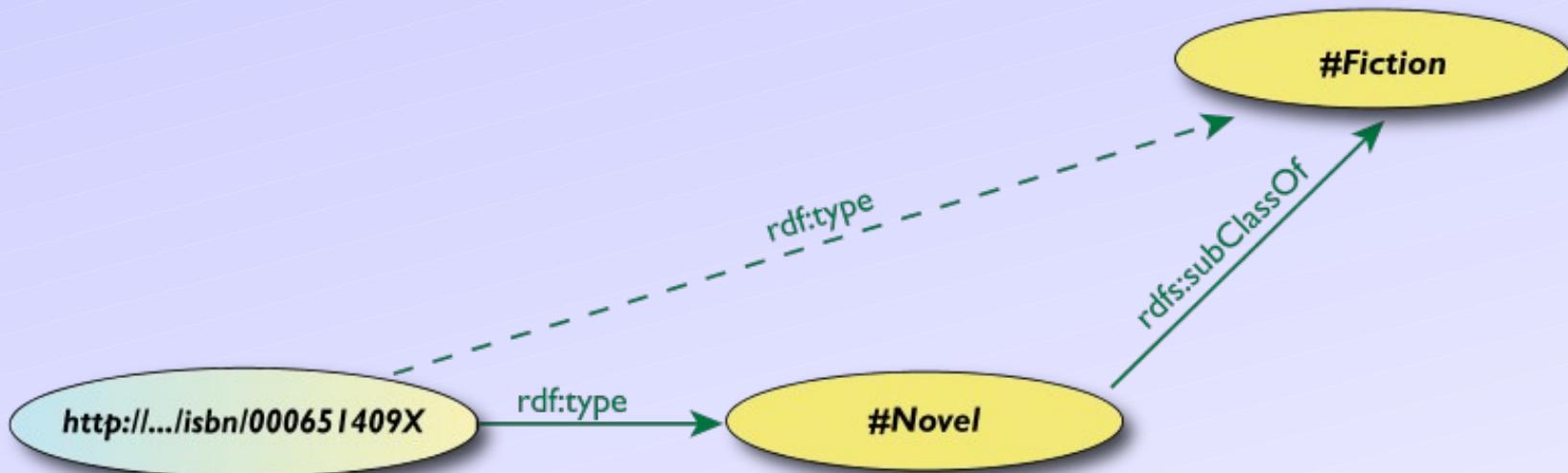
- The RDF data on a specific novel:

```
<rdf:Description rdf:about="http://.../isbn/000651409X">
  <rdf:type rdf:resource="http://.../bookSchema.rdf#Novel"/>
</rdf:Description>
```

# Further remarks on types

- A resource may belong to several classes
  - rdf:type is just a property...
    - “«The Glass Palace» is a novel, but «The Glass Palace» is also an «inventory item»...”
    - i.e., it is *not* like a datatype!
  - The type information may be very important for applications
    - e.g., it may be used for a categorization of possible nodes
    - probably the most frequently used RDF property...
      - (remember the “Person” in our example?)

# Inferred properties



```
(<http://.../isbn/000651409X> rdf:type #Fiction)
```

- is not in the original RDF data...
- ...but can be inferred from the RDFS rules
- RDFS environments return that triple, too

# Inference: let us be formal...

- The RDF Semantics document has a list of (33) entailment rules:
  - “if such and such triples are in the graph, add this and this”
  - do that recursively until the graph does not change
- The relevant rule for our example:

If:

```
uuu rdfs:subClassOf xxx .  
vvv rdf:type uuu .
```

Then add:

```
vvv rdf:type xxx .
```

# Properties

- Property is a special class (**rdf:Property**)
  - properties are also resources identified by URI-s
- There is also a possibility for a “sub-property”
  - all resources bound by the “sub” are also bound by the other
- Range and domain of properties can be specified
  - i.e., what type of resources serve as object and subject

# *Property specification serialized*

- In RDF/XML:

```
<rdf:Property rdf:ID="title">
  <rdfs:domain rdf:resource="#Fiction"/>
  <rdfs:range rdf:resource="http://...#Literal"/>
</rdf:Property>
```

- In Turtle:

```
:title
  rdf:type    rdf:Property ;
  rdfs:domain :Fiction ;
  rdfs:range  rdfs:Literal .
```

# What does this mean?

- Again, new relations can be deduced. Indeed, if

```
:title
  rdf:type    rdf:Property;
  rdfs:domain :Fiction;
  rdfs:range   rdfs:Literal.
```

```
<http://.../isbn/000651409x> :title "The Glass Palace" .
```

- then the system can infer that:

```
<http://.../isbn/000651409x> rdf:type :Fiction .
```

# Literals

- Literals may have a data type
  - floats, integers, booleans, etc, defined in XML Schemas
  - full XML fragments
- (Natural) language can also be specified

# Examples for datatypes

```
<http://.../isbn/000651409X>
:page_number "543"^^xsd:integer ;
:publ_date    "2000"^^xsd:gYear ;
:price        "6.99"^^xsd:float .
```

# *A bit of RDFS can take you far...*

- Remember the power of merge?
- We could have used, in our example:
  - **f:auteur** is a subproperty of **a:author** and vice versa (although we will see other ways to do that...)
- Of course, in some cases, more complex knowledge is necessary (see later...)

# Example: find the right experts at NASA

- Expertise locator for nearly 70,000 NASA civil servants, using RDF integration techniques over 6 or 7 geographically distributed databases, data sources, and web services...

The screenshot shows the POPS v.28.3 software interface with the following panels:

- Top Navigation:** File, Options, Bookmarks, Advanced, Help.
- Left Panel (NASA Center):** Lists facilities: ARC, DFRC, GRC, GSFC (selected), HQ, IIV, JPL, JSC, KSC, LARC, MAF, MSFC. Source: x500.
- Middle Left Panel (Project):** Lists projects: Mars Global Surveyor, Mars Odyssey 2001, Mars R&A, Mars Reconnaissance Orbiter 2005, Messenger, Minor Revital, Mission Operations, Mission Science Guest Investigator, Mission Success – Center Specific, Multi-Mission Operations, NMP Program Management and Future..., NPOESS Preparatory Project (NPP). Source: WIMS.
- Middle Right Panel (Competency):** Lists competencies: Astrobiology, Astronomy and Astrophysics, Climate Change and Variability, Earth Atmosphere, Earth Science Applications Research, Earth System Modeling, Fluid Physics, Fundamental Physics, Geophysical/Geologic Science, Geospatial Science and Technologies (selected), Icing Physics, Laser Technology. Source: CMS.
- Right Panel (People):** Lists people: Jeanne M (selected).
- Bottom Panel (Information Panel):**
  - View Different Social Network's Present in the Data:** Shows connections between Jeanne M, Michael H Grove, and Jeffrey T. The connection between Jeanne M and Michael H Grove is highlighted in pink, labeled "Skill: Earth Sciences Competency Suite" and "Project: Center Investment Accounts".
  - Social Net:** A legend indicates connection types: Same Skill and Same Department (red), Same Skill and Same Project (green), Same Skill, Project, and Facility (blue), and Am I Connected? (pink).
  - Details for Michael H Grove:** Name: Michael H Grove, Email: Michael.Grove@nasa.gov, Phone: 301-Clark and Parsia, Employer: Clark and Parsia.

# *How to get RDF Data?*

*(Microformats, GRDDL, RDFa)*

## *Simple approach*

- Write RDF/XML or Turtle “manually”
- In some cases that is necessary, but it really does not scale...

# RDF with XHTML

- Obviously, a huge source of information
- By adding some “meta” information, the same source can be reused for, eg, data integration, better mashups, etc
  - typical example: your personal information, like address, should be readable for humans and processable by machines
- Two solutions have emerged:
  - extract the structure from the page and convert the content into RDF
  - add RDF statements directly into XHTML via RDFa

## *Extract RDF*

- Use intelligent “scrapers” or “wrappers” to extract a structure (hence RDF) from a Web pages or XML files...
- ... and then generate RDF automatically (e.g., via an XSLT script)

# Formalizing the scraper approach: GRDDL

- GRDDL formalizes the scraper approach. For example:

```
<html xmlns="http://www.w3.org/1999/">
  <head profile="http://www.w3.org/2003/g/data-view">
    <title>Some Document</title>
    <link rel="transformation" href="http://.../dc-extract.xsl"/>
    <meta name="DC.Subject" content="Some subject"/>
    ...
  </head>
  ...
  <span class="date">2006-01-02</span>
  ...
</html>
```

- yields, through **dc-extract.xsl**:

```
<>
  dc:subject "Some subject";
  dc:date "2006-01-02" .
```

# GRDDL

- The transformation itself has to be provided for each set of conventions
- A more general syntax is defined for XML formats in general (e.g., via the namespace document)
  - a method to get data in other formats to RDF (e.g., XBRL)

# *Example for “structure”: microformats*

- Not a Semantic Web specification, originally
  - there is a separate microformat community
- Approach: re-use (X)HTML attributes and elements to add “meta” information
  - typically @abbr, @class, @title, ...
  - different community agreements for different applications

# RDFa

- RDFa extends (X)HTML a bit by:
  - defining general attributes to add metadata to any elements
  - provides an almost complete “serialization” of RDF in XHTML
- It is a bit like the microformats/GRDDL approach but fully generic

# RDFa example

- For example:

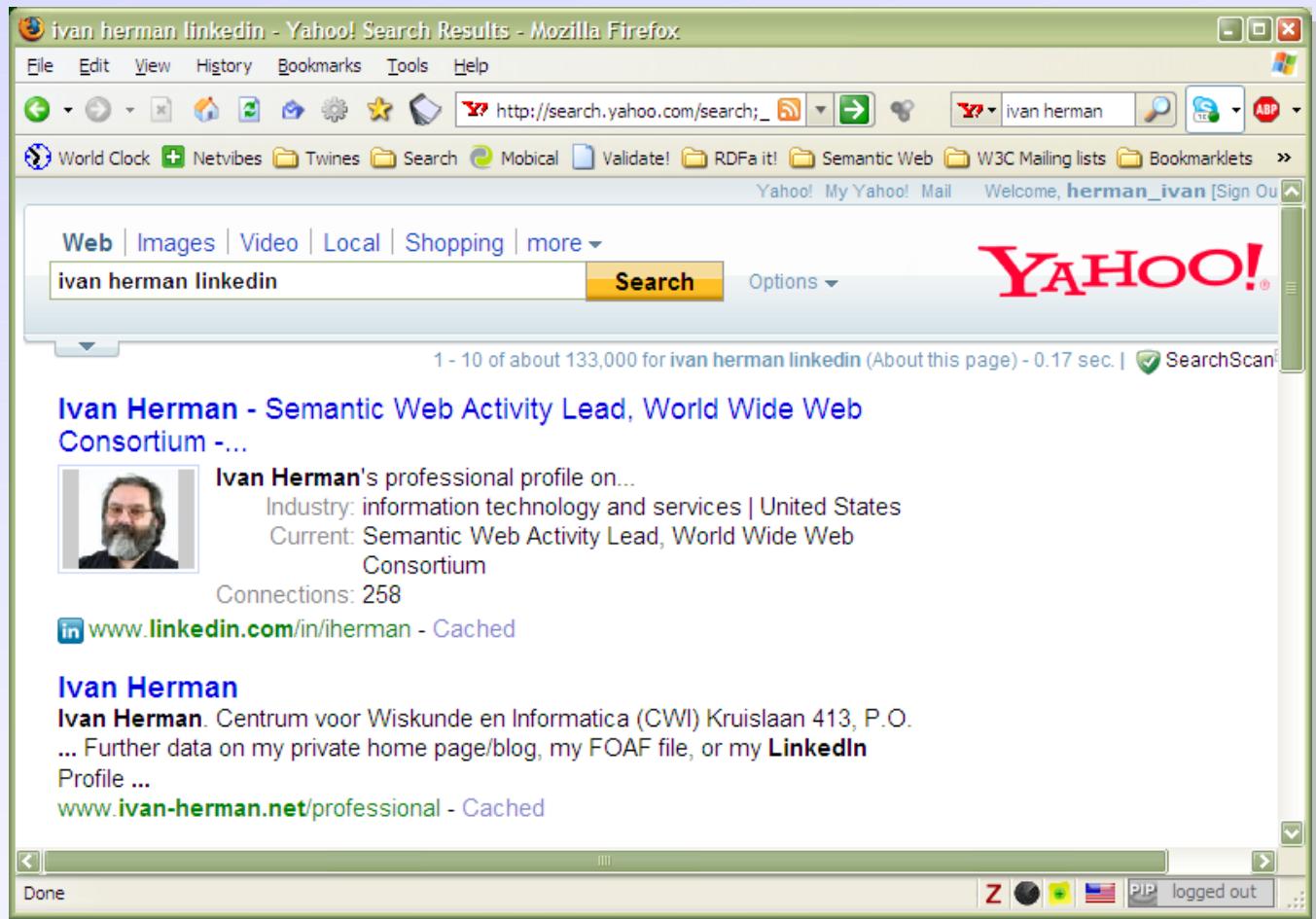
```
<div about="http://uri.to.newsitem">
  <span property="dc:date">March 23, 2004</span>
  <span property="dc:title">Rollers hit casino for £1.3m</span>
  By <span property="dc:creator">Steve Bird</span>. See
  <a href="http://www.a.b.c/d.avi" rel="dc:type:MovingImage">
    also video footage</a>...
</div>
```

- yields, through an RDFa processor:

```
<http://uri.to.newsitem>
  dc:date          "March 23, 2004";
  dc:title         "Rollers hit casino for £1.3m";
  dc:creator       "Steve Bird";
  dc:type:MovingImage <http://www.a.b.c/d.avi>.
```

# Example: Yahoo's SearchMonkey

- Search based results may be customized via small applications
- Metadata in pages (in RDFa, microformats etc) are reused



# Example: RDFa data by the London Gazette

Search Results - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://www.gazettes-online.co.uk/ViewGazetteDocument.aspx?atdocid=5437822&Ge

Saturday, November 20,

The London Gazette

Change edition: Edinburgh / Belfast

Home

About the Gazette

Browse

Search Tools

My Account

My Notices

Services

Placing a Notice

Help

Search archive Enter keyword  Advanced Search

## Search Results

Results 0 of 14 gazette documents

[Back to results](#)

Documents: [Previous](#) [10](#) [11](#) [12](#) [13](#) [14](#) [Next](#)

Date: 31 October 2008 Issue Number: 58870 Page number: 16858

Publication Date: *Friday, 31 October 2008*

Notice Code: *1901*

Water Resources

*Environment Agency*

RDFa

Done

# Example: RDFa data by the London Gazette

http://www.w3.org/2007/08/pyRdfa/extract?uri=http://www.gazettes-online.co.uk/ViewGazetteDocument.aspx?atdocid=5437822&GeoType... □ X

File Edit View History Bookmarks Tools Help

Search Results http://www.w3.org/2007/08/pyRdfa/e... □

http://www.w3.org/2007/08/pyRdfa/extract?uri=http%3A//www.gazettes-online.co.uk/Vie... □

Wikia Search ABP

GeoType=London&categorydocids=144&lastissuecount=10

http://www.gazettes-online.co.uk/ViewGazetteDocument.aspx?atdocid=5437822&GeoType=London&categorydocids=144&lastissuecount=10	stylesheet	http://www.gazettes-online.co.uk/Styles/gazettes.css
London Gazette: Issue dated 31 October 2008: Notice 650554	Creator	TSO (The Stationery Office)
	Identifier	http://www.london-gazette.co.uk/issues/2008-10-31/notices/650554
	Language	Member Of ISO 639-2
	Publisher	value eng
		TSO (The Stationery Office), St Crispins, Duke Street, Norwich, NR3 1PD, 01603 622211, customer.services@tso.co.uk
	Subject	Member Of IPSV
		value Water Resources
	Title	London Gazette: Issue dated 31 October 2008: Notice 650554
	Date Issued	2008-10-31
	Date Modified	2008-08-20
	Administrator	Grant Wilson
	Authority	Environment Agency
	Category Code	1901
	Notice Number	650554
	Publication Date	2008-10-31
	is In Issue	http://www.london-gazette.co.uk/issues/2008-10-31
	type	Water Resources Notice
http://www.london-gazette.co.uk/issues/2008-10-31	Issue Number	58870
	Publication Date	2008-10-31
Environment Agency	is Known As	Environment Agency
	type	Authority
Grant Wilson	Forename	Grant
	Surname	Wilson
	type	Person

Done

# *Bridge to relational databases*

- Data on the Web are mostly stored in databases
- “Bridges” are being defined:
  - a layer between RDF and the relational data
    - RDB tables are “mapped” to RDF graphs, possibly on the fly
    - different mapping approaches are being used
  - a number RDB systems offer this facility already (eg, Oracle, OpenLink, ...)
- A survey on mapping techniques has been published at W3C
- W3C plans to engage in a standardization work in this area

# *Linking Data*

# Linking Open Data Project

- Goal: “expose” open datasets in RDF
- Set *RDF links among the data items* from different datasets
- Set up query endpoints
- Altogether billions of triples, millions of links...



# Example data source: DBpedia

- DBpedia is a community effort to
  - extract structured (“infobox”) information from Wikipedia
  - provide a query endpoint to the dataset
  - interlink the DBpedia dataset with other datasets on the Web



UNIVERSITÄT LEIPZIG



# Extracting Wikipedia structured data

Amsterdam	
	
The Keizersgracht at dusk	
Location of Amsterdam	
Coordinates:	52°22'23"N 4°53'32"E
Country	Netherlands
Province	North Holland
Government	
- Type	Municipality
- Mayor	Job Cohen <sup>[1]</sup> (PvdA)
- Aldermen	Lodewijk Asscher Carolien Gehrels Tjeerd Herrema Maarten van Poelgeest Marijke Vos Erik Gerritsen
- Secretary	
Area <sup>[2][3]</sup>	
- City	219 km <sup>2</sup> (84.6 sq mi)
- Land	166 km <sup>2</sup> (64.1 sq mi)
- Water	53 km <sup>2</sup> (20.5 sq mi)
- Urban	1,003 km <sup>2</sup> (387.3 sq mi)
- Metro	1,815 km <sup>2</sup> (700.8 sq mi)
Elevation <sup>[4]</sup>	2 m (7 ft)
Population <sup>(1 October 2008)[5][6]</sup>	
- City	755,269
- Density	4,459/km <sup>2</sup> (11,548.8/sq mi)
- Urban	1,364,422
- Metro	2,158,372
- Demonym	Amsterdamer
Time zone	CET (UTC+1)
- Summer (DST)	CEST (UTC+2)
Postcodes	1011 – 1109
Area code(s)	020
Website: <a href="http://www.amsterdam.nl">www.amsterdam.nl</a>	

@prefix dbpedia <<http://dbpedia.org/resource/>>.

@prefix dbterm <<http://dbpedia.org/property/>>.

dbpedia:Amsterdam

dbterm:officialName "Amsterdam" ;

dbterm:longd "4" ;

dbterm:longm "53" ;

dbterm:longs "32" ;

...

dbterm:leaderTitle "Mayor" ;

dbterm:leaderName dbpedia:Job\_Cohen ;

...

dbterm:areaTotalKm "219" ;

...

dbpedia:ABN\_AMRO

dbterm:location dbpedia:Amsterdam ;

...

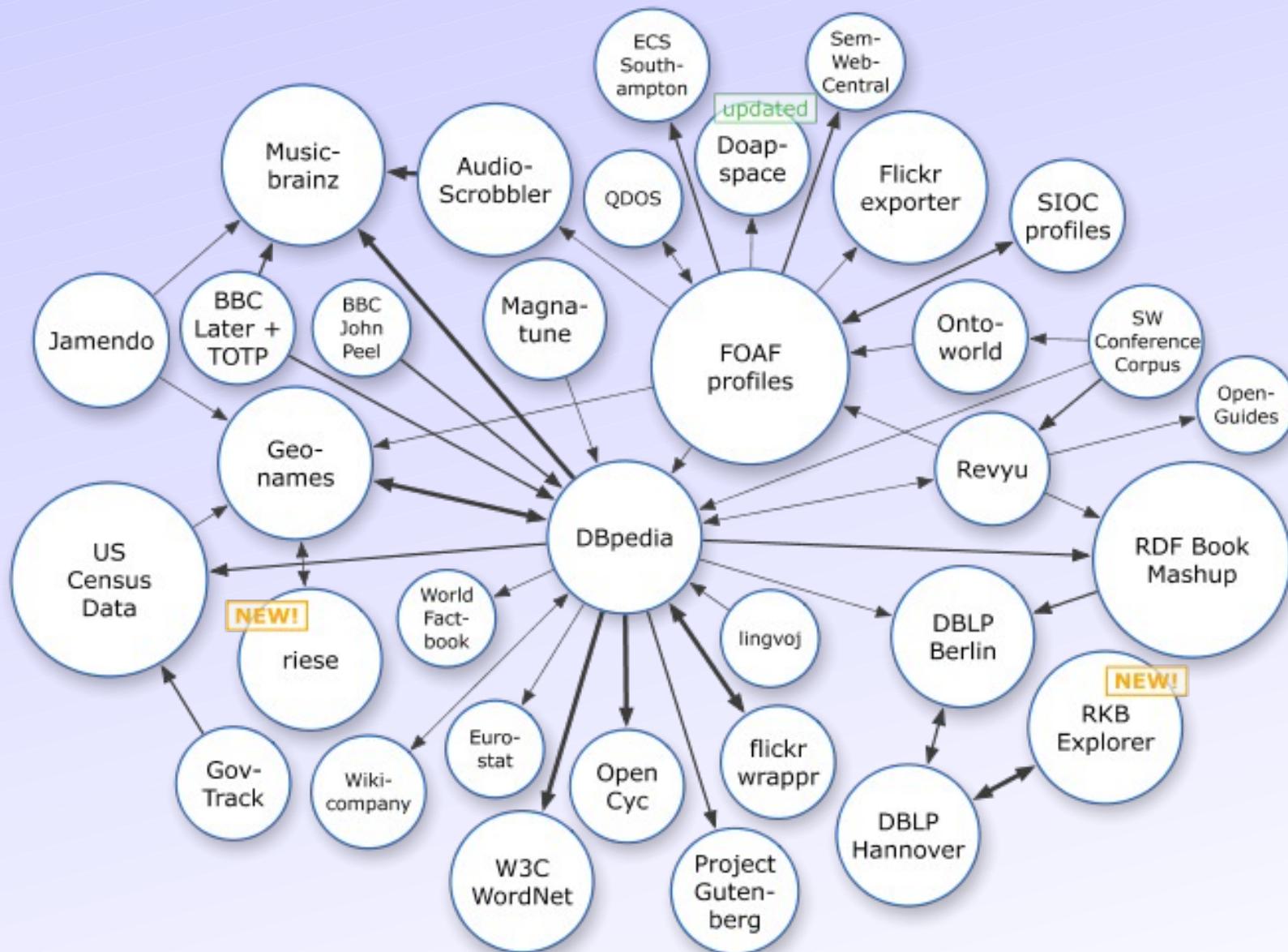
# Automatic links among open datasets

```
<http://dbpedia.org/resource/Amsterdam>
owl:sameAs <http://rdf.freebase.com/ns/...> ;
owl:sameAs <http://sws.geonames.org/2759793> ;
...
```

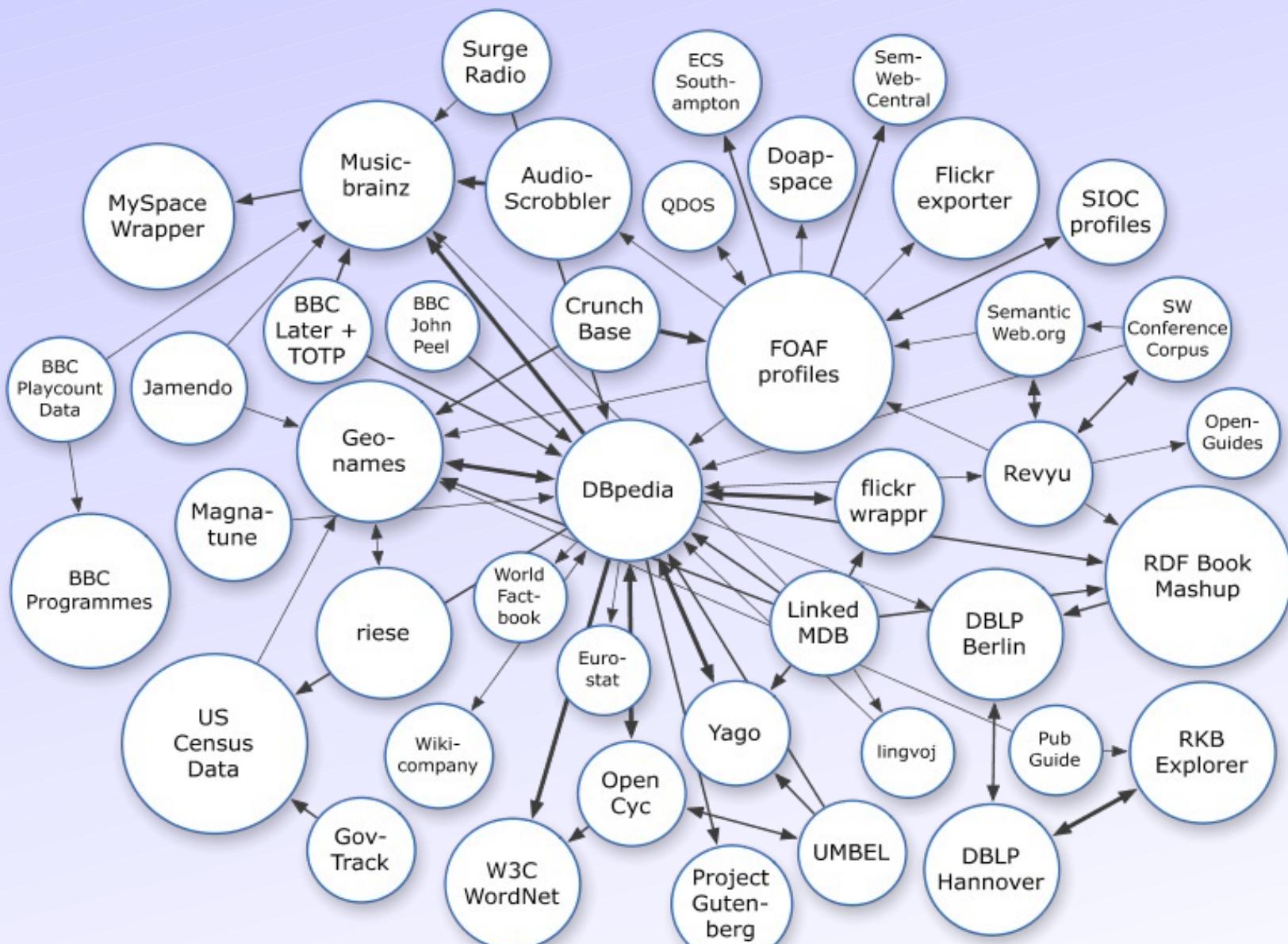
```
<http://sws.geonames.org/2759793>
owl:sameAs <http://dbpedia.org/resource/Amsterdam>
wgs84_pos:lat "52.3666667" ;
wgs84_pos:long "4.8833333" ;
geo:inCountry <http://www.geonames.org/countries/#NL> ;
...
```

Processors can switch automatically from one to the other...

# The LOD “cloud”, March 2008

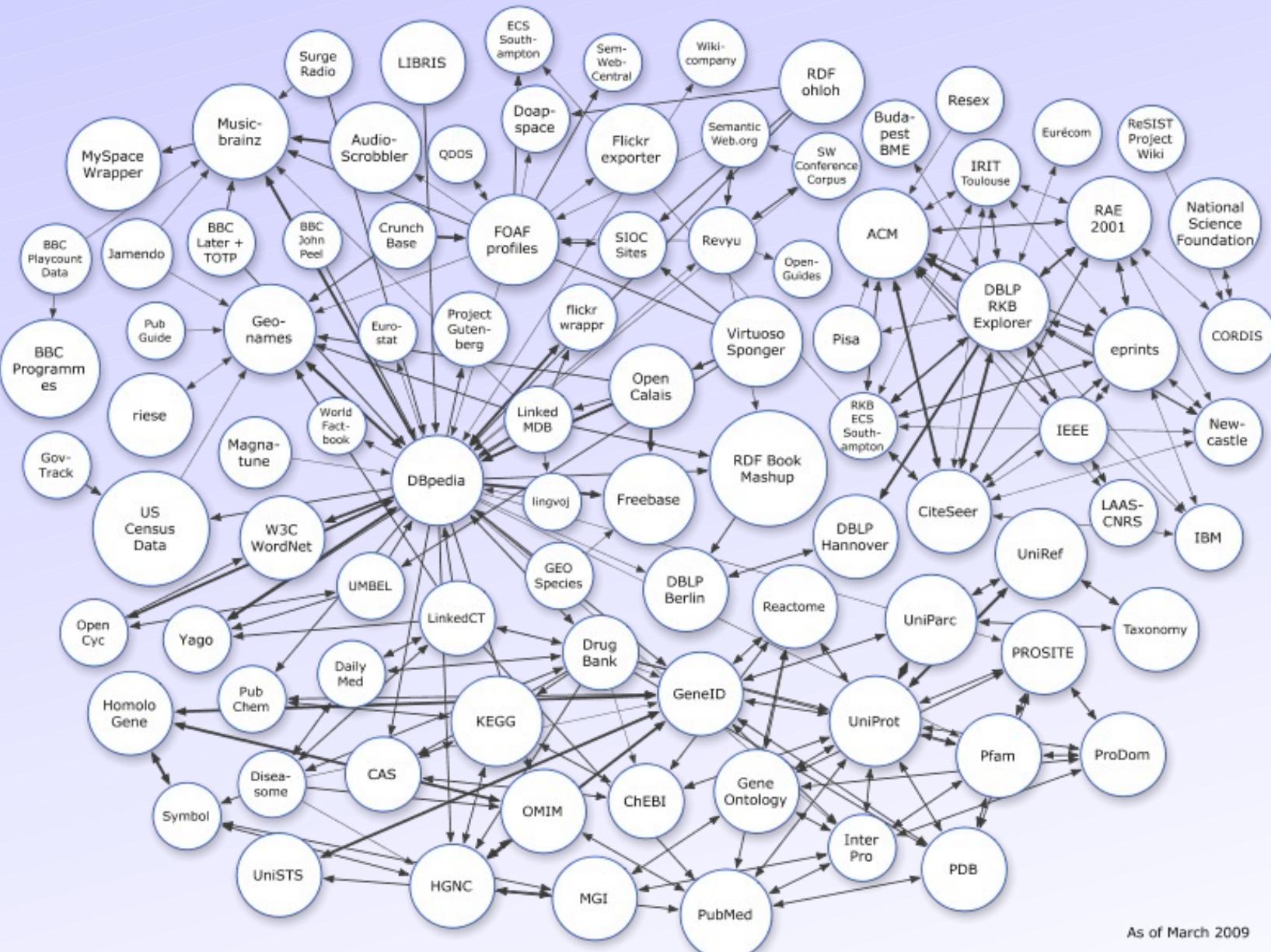


# *The LOD “cloud”, September 2008*



As of September 2008

# The LOD “cloud”, March 2009



# Example: mapping application on an iPhone



# Example: mapping application on an iPhone



# *Query RDF Data (SPARQL)*

# RDF data access

- How do I query the RDF data?
  - e.g., how do I get to the DBpedia data?

# Querying RDF graphs

- Remember the Jena idiom:

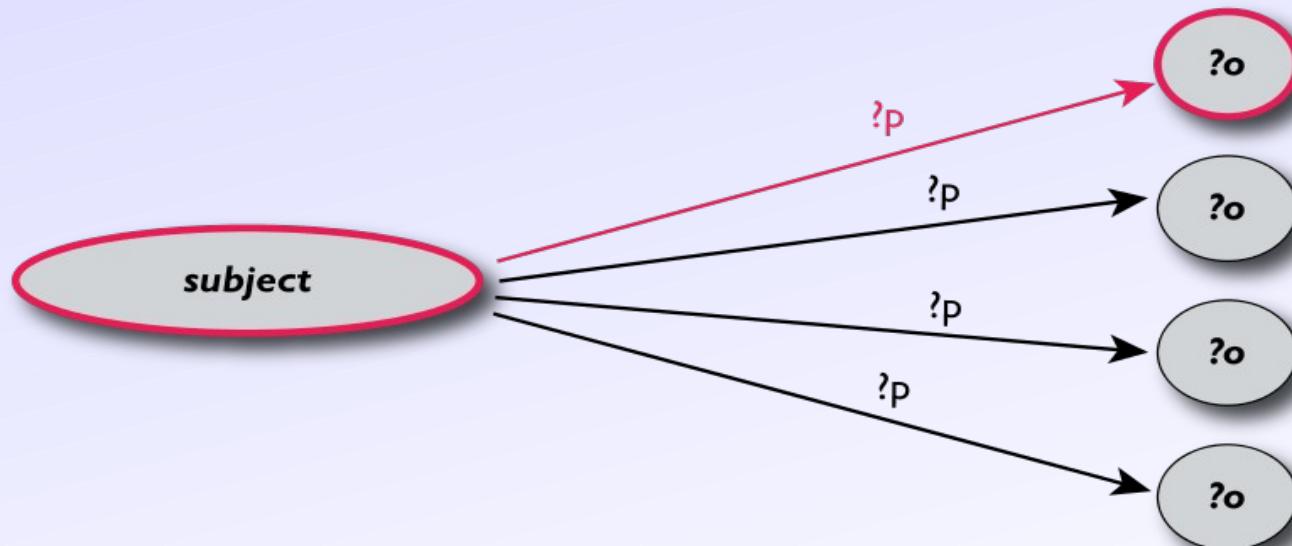
```
StmtIterator iter=model.listStatements(subject,null,null);  
while(iter.hasNext()) {  
    st = iter.next();  
    p = st.getProperty(); o = st.getObject();  
    do_something(p,o);
```

- In practice, more complex queries into the RDF data are necessary
  - something like: “give me the (a,b) pair of resources, for which there is an x such that (x parent a) and (b brother x) holds” (ie, return the uncles)
  - these rules may become quite complex
- The goal of SPARQL (Query Language for RDF)

# Analyse the Jena example

```
StmtIterator iter=model.listStatements(subject,null,null);  
while(iter.hasNext()) {  
    st = iter.next();  
    p = st.getProperty(); o = st.getObject();  
    do_something(p,o);
```

- The  $(\text{subject}, ?p, ?o)$  is a *pattern* for what we are looking for (with  $?p$  and  $?o$  as “unknowns”)



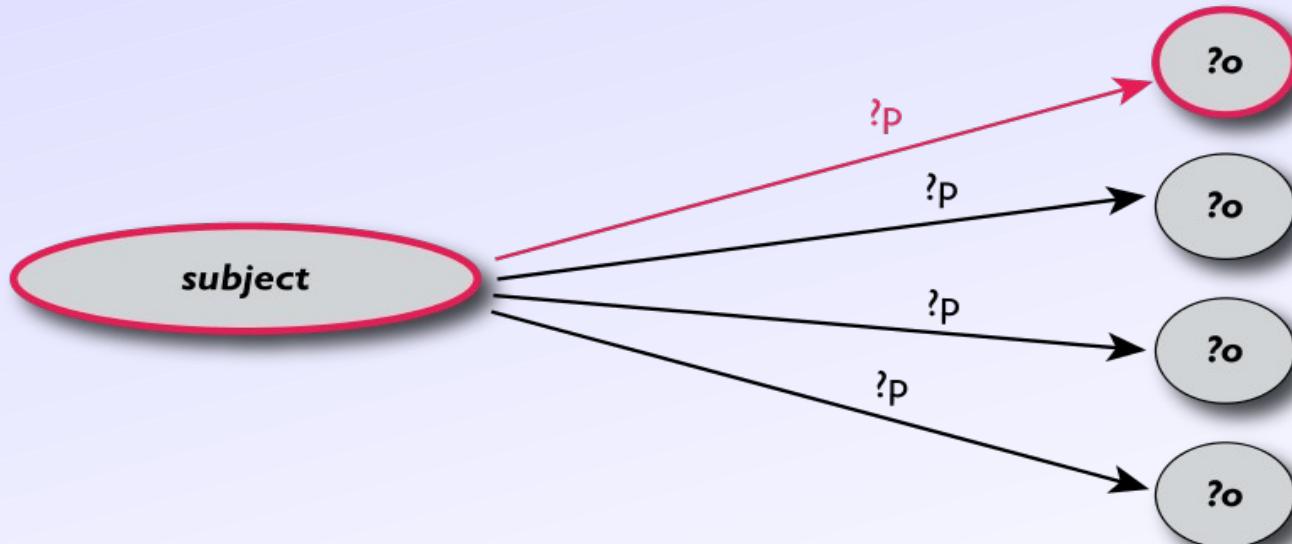
# *General: graph patterns*

- The fundamental idea: use graph patterns
  - the pattern contains unbound symbols
  - by binding the symbols, subgraphs of the RDF graph are selected
  - if there is such a selection, the query returns bound resources

# Our Jena example in SPARQL

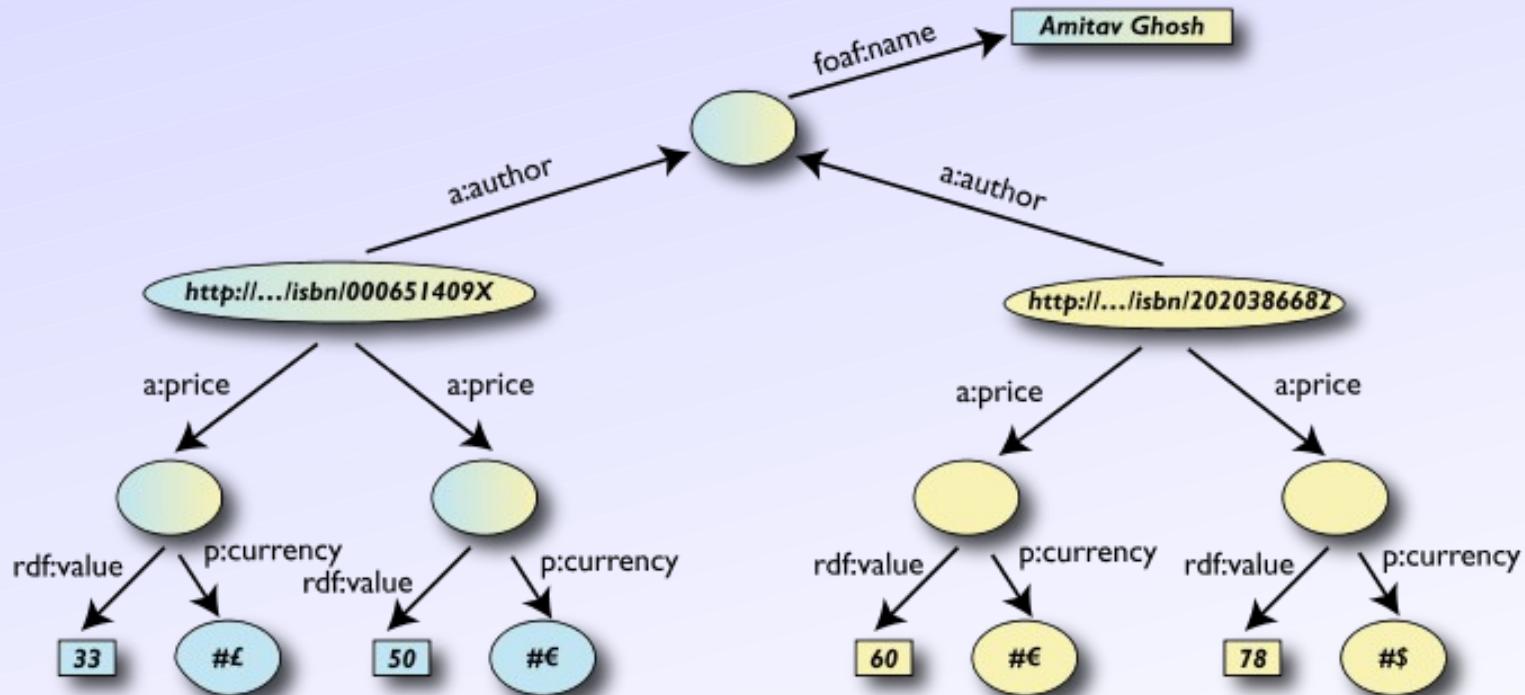
```
SELECT ?p ?o  
WHERE {subject ?p ?o}
```

- The triples in WHERE define the graph pattern, with ?p and ?o “unbound” symbols
- The query returns all p,o pairs



# Simple SPARQL example

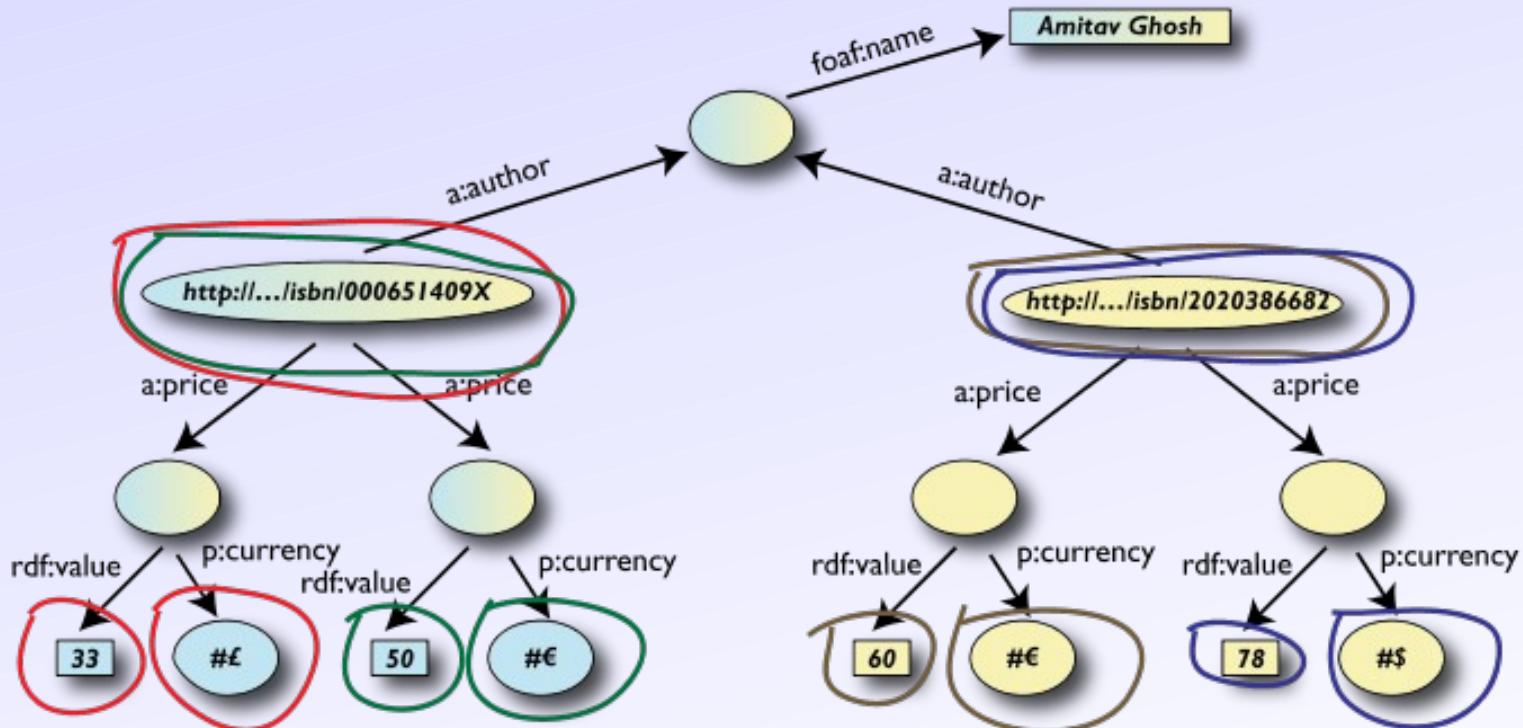
```
SELECT ?isbn ?price ?currency # note: not ?x!
WHERE {?isbn a:price ?x. ?x rdf:value ?price. ?x p:currency ?currency.}
```



# Simple SPARQL example

```
SELECT ?isbn ?price ?currency # note: not ?x!
WHERE {?isbn a:price ?x. ?x rdf:value ?price. ?x p:currency ?currency.}
```

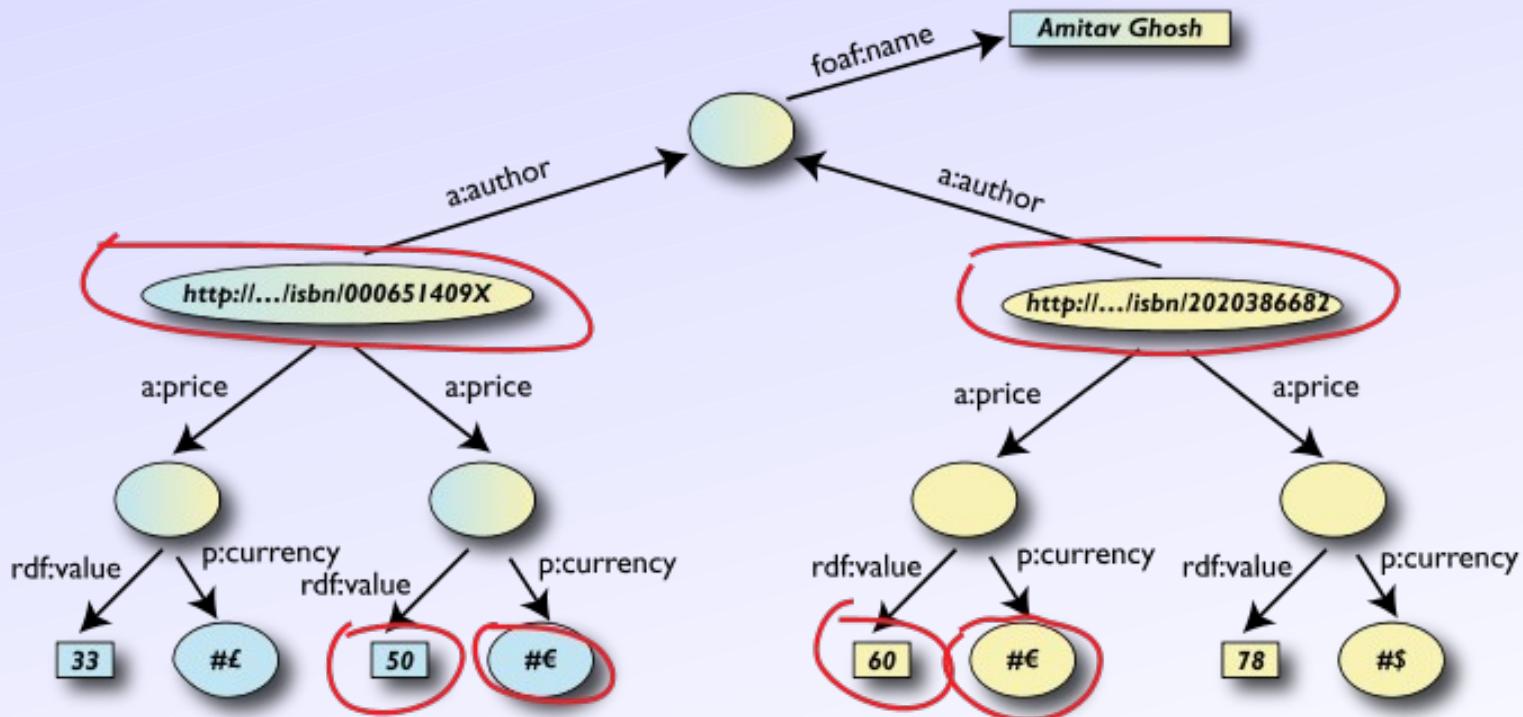
- Returns:  
[[<..49X>,33,£], [<..49X>,50,€], [<..6682>,60,€],  
[<..6682>,78,\$]]



# Pattern constraints

```
SELECT ?isbn ?price ?currency # note: not ?x!
WHERE { ?isbn a:price ?x. ?x rdf:value ?price. ?x p:currency ?currency .
        FILTER(?currency == €) }
```

- Returns: [[<..409X>,50,€], [<..6682>,60,€]]



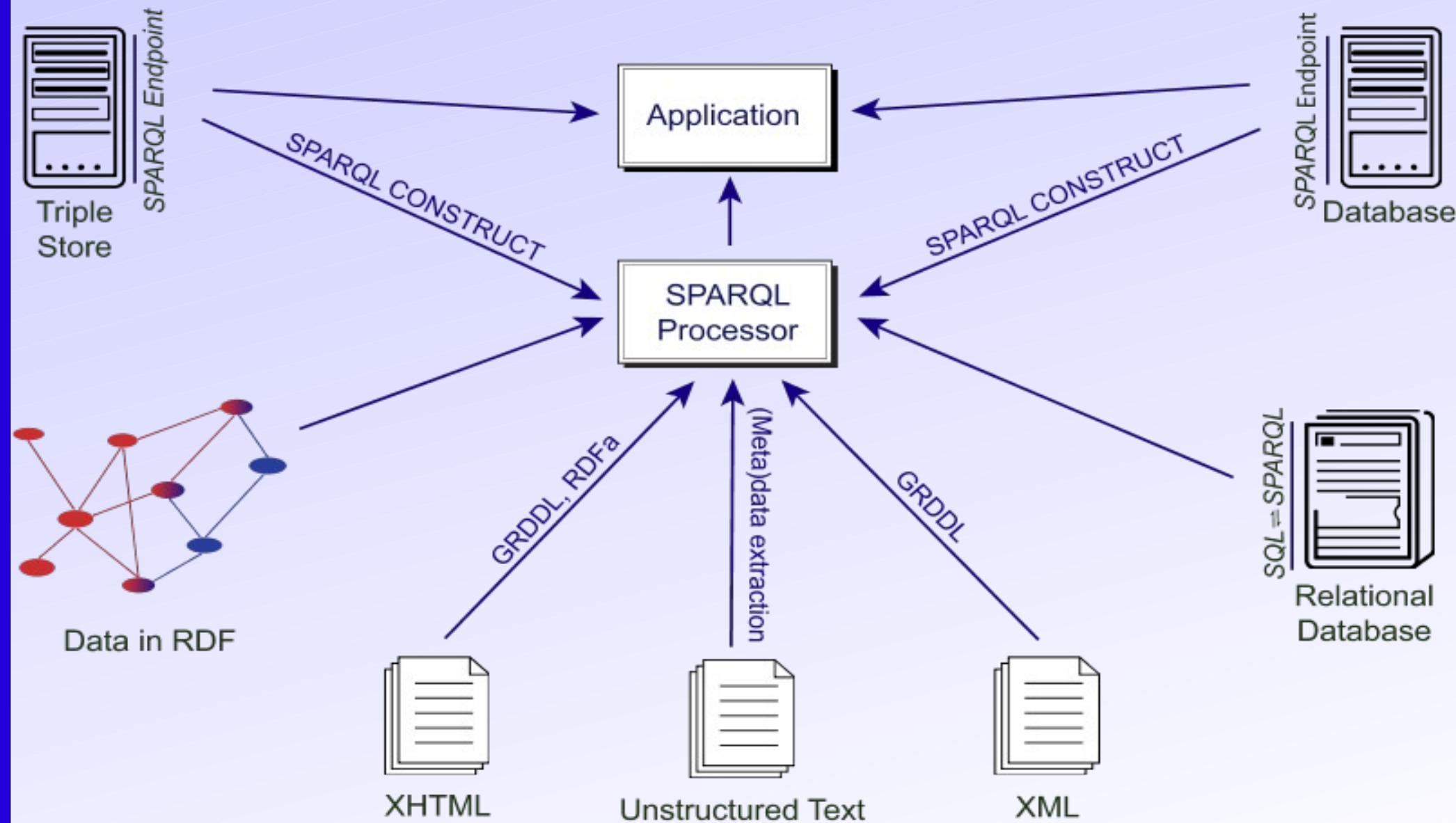
# Other SPARQL features

- Limit the number of returned results; remove duplicates, sort them, ...
- Optional branches in the query
- Specify several data sources (via URI-s) within the query (essentially, a merge!)
- Construct a graph combining a separate pattern and the query results
- Use datatypes and/or language tags when matching a pattern

# *SPARQL usage in practice*

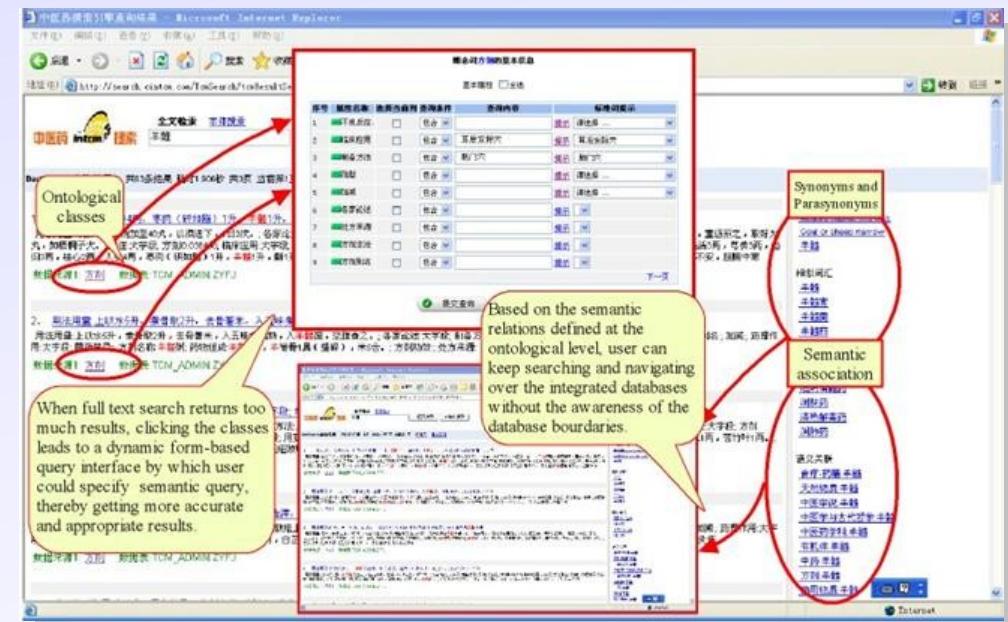
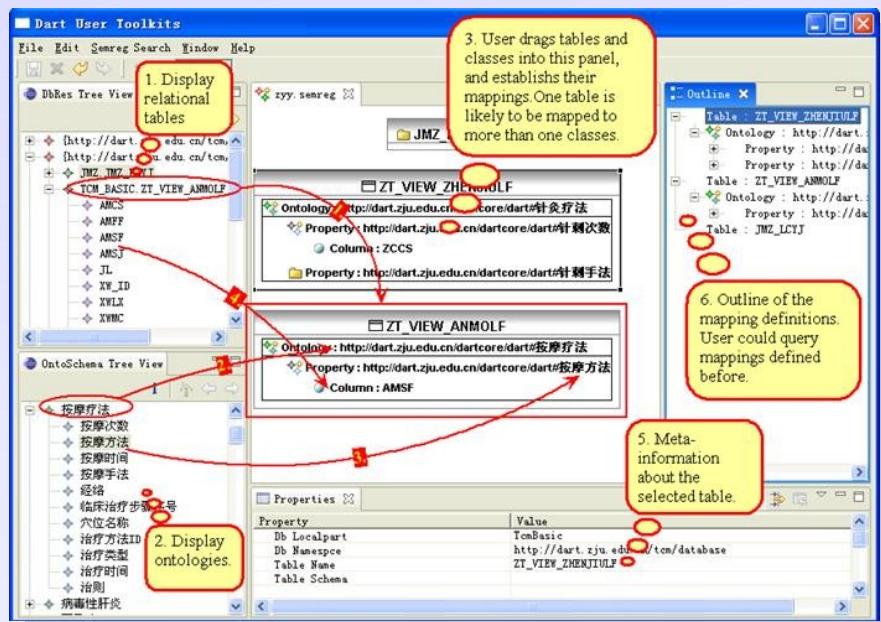
- SPARQL is usually used over the network
  - separate documents define the protocol and the result format
  - SPARQL Protocol for RDF with HTTP and SOAP bindings
  - SPARQL results in XML or JSON formats
- Big datasets usually offer “SPARQL endpoints” using this protocol
  - typical example: SPARQL endpoint to DBpedia

# SPARQL as a unifying point



# Example: integrate Chinese medical data

- Integration of a large number of TCM databases
  - around 80 databases, around 200,000 records each
- A visual tool to map databases to the semantic layer using a specialized ontology
- Form based query interface for end users



# *Ontologies (OWL)*

# Ontologies

- RDFS is useful, but does not solve all possible requirements
- Complex applications may want more possibilities:
  - characterization of properties
  - identification of objects with different URI-s
  - disjointness or equivalence of classes
  - construct classes, not only name them
  - can a program reason about some terms? E.g.:
    - “if «Person» resources «A» and «B» have the same «**foaf:email**» property, then «A» and «B» are identical”
  - etc.

# Ontologies (cont.)

- The term ontologies is used in this respect:

“defines the concepts and relationships used to describe and represent an area of knowledge”

- RDFS can be considered as a simple ontology language
- Languages should be a compromise between
  - rich semantics for meaningful applications
  - feasibility, implementability

# *Web Ontology Language = OWL*

- OWL is an extra layer, a bit like RDF Schemas
  - own namespace, own terms
  - it relies on RDF Schemas
- It is a separate recommendation
  - actually... there is a 2004 version of OWL (“OWL 1”)
  - and there is an update (“OWL 2”) that should be finalized in 2009
  - you will surely hear about it at the conference...

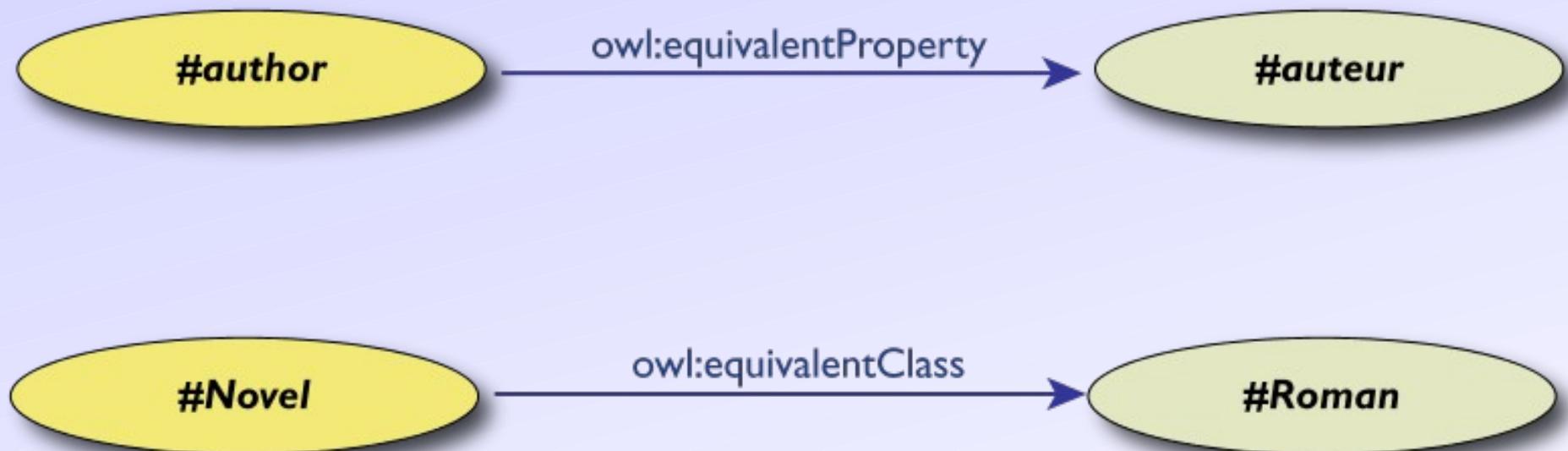
# *OWL is complex...*

- OWL is a large set of additional terms
- We will not cover the whole thing here...

# Term equivalences

- For classes:
  - **owl:equivalentClass**: two classes have the same individuals
  - **owl:disjointWith**: no individuals in common
- For properties:
  - **owl:equivalentProperty**
    - remember the **a:author** vs. **f:auteur**
  - **owl:propertyDisjointWith**
- For individuals:
  - **owl:sameAs**: two URIs refer to the same concept (“individual”)
  - **owl:differentFrom**: negation of **owl:sameAs**

# Connecting to French...



## *Typical usage of owl:sameAs*

- Linking our example of Amsterdam from one data set (DBpedia) to the other (Geonames):

```
<http://dbpedia.org/resource/Amsterdam>
owl:sameAs <http://sws.geonames.org/2759793>;
```

- This is the main mechanism of “Linking” in the Linking Open Data project

# Property characterization

- In OWL, one can characterize the behaviour of properties (symmetric, transitive, functional, inverse functional...)
- One property may be the inverse of another
- OWL also separates *data* and *object* properties
  - “datatype property” means that its range are typed literals

# What this means is...

- If the following holds in our triples:

```
:email rdf:type owl:InverseFunctionalProperty.  
<A> :email "mailto:a@b.c".  
<B> :email "mailto:a@b.c".
```

then, processed through OWL, the following holds, too:

```
<A> owl:sameAs <B>.
```

- I.e., *new relationships* were discovered again (beyond what RDFS could do)

# Classes in OWL

- In RDFS, you can subclass existing classes... that's all
- In OWL, you can construct classes from existing ones:
  - enumerate its content
  - through intersection, union, complement
  - Etc

# Classes in OWL (cont)

- OWL makes a stronger conceptual distinction between classes and individuals
  - there is a separate term for `owl:Class`, to make the difference (a specialization of the RDFS class)
  - individuals are separated into a special class called `owl:Thing`
- Eg, a precise classification would be:

```
ex:Person rdf:type owl:Class.
```

```
<uri-for-Amitav-Ghosh>
  rdf:type owl:Thing;
  rdf:type owl:Person .
```

# Classes contents can be enumerated

```
:£ rdf:type owl:Thing.  
:€ rdf:type owl:Thing.  
:$ rdf:type owl:Thing.  
:Currency  
    rdf:type owl:Class;  
    owl:oneOf (:€ :£ :$).
```

- I.e., the class consists of exactly of those individuals

# Union of classes can be defined

```
:Novel          rdf:type owl:Class.  
:Short_Story   rdf:type owl:Class.  
:Poetry        rdf:type owl:Class.  
:Literature    rdf:type owl:Class;  
               owl:unionOf (:Novel :Short_Story :Poetry).
```

- Other possibilities: `complementOf`, `intersectionOf`, ...

# For example...

If:

```
:Novel          rdf:type owl:Class.  
:Short_Story    rdf:type owl:Class.  
:Poetry         rdf:type owl:Class.  
:Literature     rdf:type owl:Class;  
               owl:unionOf (:Novel :Short_Story :Poetry).  
  
<myWork> rdf:type :Novel .
```

then the following holds, too:

```
<myWork> rdf:type :Literature .
```

# *It can be a bit more complicated...*

If:

```
:Novel          rdf:type owl:Class.  
:Short_Story   rdf:type owl:Class.  
:Poetry        rdf:type owl:Class.  
:Literature    rdf:type owlClass;  
               owl:unionOf (:Novel :Short_Story :Poetry).
```

```
fr:Roman owl:equivalentClass :Novel .
```

```
<myWork> rdf:type fr:Roman .
```

then, through the *combination* of different terms,  
the following still holds:

```
<myWork> rdf:type :Literature .
```

## *What we have so far...*

- The OWL features listed so far are already fairly powerful
- E.g., various databases can be linked via **owl:sameAs**, functional or inverse functional properties, etc.
- Many inferred relationships can be found using a traditional rule engine

## ***However... that may not be enough***

- Very large vocabularies might require even more complex features
  - typical usage example: definition of all concepts in a health care environment
  - a major issue: the way classes (i.e., “concepts”) are defined
- OWL includes those extra features but... the inference engines become (much) more complex 😞

# Property value restrictions

- Classes are created by restricting the property values on its individuals
- For example: how would I characterize a “listed price”?
  - it is a price (which may be a general term), but one that is given in one of the “allowed” currencies (say, €, £, or \$)
  - more formally:
    - the value of “**p:currency**”, when applied to a resource on listed price, must be of one of those values...
    - ...thereby defining the class of “listed price”

# *Restrictions formally*

- Defines a class of type **owl:Restriction** with a
  - reference to the property that is constrained
  - definition of the constraint itself
- One can, e.g., subclass from this node when defining a particular class

```
:Listed_Price rdfs:subClassOf [  
    rdf:type owl:Restriction;  
    owl:onProperty p:currency;  
    owl:allValuesFrom :Currency.  
].
```

# Possible usage...

If:

```
:Listed_Price rdfs:subClassOf [  
    rdf:type owl:Restriction;  
    owl:onProperty p:currency;  
    owl:allValuesFrom :Currency.  
].  
  
:price rdf:type :Listed_Price .  
  
:price p:currency <something> .
```

then the following holds:

```
<something> rdf:type :Currency .
```

# Other restrictions

- **allValuesFrom** could be replaced by:
  - **someValuesFrom**
    - e.g., I could have said: there should be a price given in at least one of those currencies
    - **hasValue**, when restricted to one specific value
- Cardinality restrictions: instead of looking at the values of properties, their number is considered
  - eg, a specific property should occur exactly once

# *But: OWL is hard!*

- The combination of class constructions with various restrictions is extremely powerful
- What we have so far follows the same logic as before
  - extend the basic RDF and RDFS possibilities with new features
  - define their semantics, ie, what they “mean” in terms of relationships
  - expect to infer new relationships based on those
- However... a full inference procedure is hard 🤯
  - not implementable with simple rule engines, for example

# OWL “species”

- OWL species comes to the fore:
  - restricting which terms can be used and under what circumstances (restrictions)
  - if one abides to those restrictions, then simpler inference engines can be used
- They reflect compromises: expressibility vs. implementability

# *OWL Full*

- No constraints on any of the constructs
  - `owl:Class` is just syntactic sugar for `rdfs:Class`
  - `owl:Thing` is equivalent to `rdfs:Resource`
  - this means that:
    - Class can also be an individual, a URI can denote a property as well as a Class
      - e.g., it is possible to talk about class of classes, apply properties on them
      - etc
    - etc.
- Extension of RDFS in all respects
- But: no system may exist that infers everything one might expect

# *OWL Full usage*

- Nevertheless OWL Full is essential
  - it gives a generic framework to express many things
  - some application just need to express and interchange terms (with possible scruffiness)
- Applications may control what terms are used and how
  - in fact, they may define their own sub-language via, eg, a vocabulary
    - thereby ensuring a manageable inference procedure

# OWL DL

- A number of restrictions are defined
  - classes, individuals, object and datatype properties, etc, are fairly strictly separated
  - object properties must be used with individuals
    - i.e., properties are really used to create relationships between individuals
    - no characterization of datatype properties
    - ...
- But: well known inference algorithms exist!

# Examples for restrictions

- The following is not “legal” OWL DL:

```
<q> rdf:type <A>.           # A is a class, q is an individual  
  
<r> rdf:type <q>.           # error: q cannot be used for a class, too  
  
<A> ex:something <B>.       # error: properties are for individuals only  
  
<q> ex:something <s>.         # error: same property cannot be used as  
<p> ex:something "54".       #   object and datatype property
```

# *OWL DL usage*

- Abiding to the restrictions means that very large ontologies can be developed that require precise procedures
  - eg, in the medical domain, biological research, energy industry, financial services (eg, XBRL), etc
  - the number of classes and properties described this way can go up to the many thousands
- OWL DL has become a language of choice to define and manage formal ontologies in general
  - even if their usage is not necessarily on the Web

# ***OWL 2 defines further species a.k.a. “profiles”***

- Further restrictions on how terms can be used and what inferences can be expected
  - Classification and instance queries in polynomial time: *OWL-EL*
  - Implementable on top of conventional relational database engines: *OWL-QL*
  - Implementable on top of traditional rule engines: *OWL-RL*

# Ontology development

- The hard work is to create the ontologies
  - requires a good knowledge of the area to be described
  - some communities have good expertise already (e.g., librarians)
  - OWL is just a tool to formalize ontologies
  - large scale ontologies are often developed in a community process
- Ontologies should be shared and reused
  - can be via the simple namespace mechanisms...
  - ...or via explicit import

# Must I use large ontologies?

- NO!!!
- Many applications are possible with RDFS and a just a little bit of OWL
  - a few terms, whose meaning is defined in OWL, and that application can handle directly
  - OWL RL is a step to create such a generic OWL level
- Big ontologies can be expensive (both in time and money); use them only when really necessary!

# *Ontologies examples*

- eClassOwl: eBusiness ontology for products and services, 75,000 classes and 5,500 properties
- National Cancer Institute's ontology: about 58,000 classes
- Open Biomedical Ontologies Foundry: a collection of ontologies, including the Gene Ontology to describe gene and gene product attributes in any organism or protein sequence and annotation terminology and data (UniProt)
- BioPAX: for biological pathway data

# Example: improved search via ontology

- Search results are re-ranked using ontologies
- Related terms are highlighted, usable for further search

The screenshot shows the GoPubMed interface in Mozilla Firefox. The search term 'tinnitus' is entered in the search bar, resulting in 1,000 articles. The left sidebar displays a tree view of 'Top categories' under 'what', with a red oval highlighting the 'Diseases [985]' section. The main content area lists four search results, each with a thumbnail, author information, PMID, title, and abstract. A blue arrow points from the third result back to the 'Diseases' category in the sidebar.

Result Number	Title	Author(s)	PMID	Abstract Excerpt
5	Pros and cons of tinnitus retraining therapy.	Hatanaka A et al., Acta Otolaryngol, 128 (4): 365-8, 2008	18368566	A significant reduction in the <b>tinnitus</b> Handicap Inventory (THI) was obtained as early as 1 month after implementation of <b>tinnitus</b> retraining therapy (TRT).
1	Gabapentin effectiveness on the sensation of subjective idiopathic tinnitus : a pilot study.	Bakhshaei M et al., Eur Arch Otorhinolaryngol, 2007	17960408	Pure-tone audiograms, laboratory test and personal histories were used to exclude any particular etiology of <b>tinnitus</b> .
3	Algorithm for evaluation of pulsatile tinnitus.	Mattox DE et al., Acta Otolaryngol, 128 (4): 427-31, 2008	18368578	Among patients with venous <b>tinnitus</b> , sigmoid sinus diverticulum was the most common finding.
4	Functional imaging of unilateral tinnitus using fMRI.	Lanting CP et al., Acta Otolaryngol, 128 (4): 415-21, 2008	18368576	The response to sound in the inferior colliculus was elevated in <b>tinnitus</b> patients compared with controls without <b>tinnitus</b> .

# Example: improved search via ontology

- Same dataset, different ontology
  - (ontology is on non-animal experimentation)

Screenshot of the Go3R - Mozilla Firefox interface showing search results for "tinnitus".

The search bar shows "tinnitus" and the result count is "1,000 articles".

The left sidebar displays a tree view of categories under "what", with several nodes circled in red:

- 3R Relevance Filters (Beta)
- Top categories
  - Diseases & Symptoms [601]
    - Tinnitus [547]
    - Hearing Loss [248]
    - Vertigo [98]
    - Disease [118]
    - Hearing Loss, Sensorineural [95]
  - more
  - Methodology [408]
  - Life Sciences [503]
  - Body Systems & Structures [401]
    - Bioethics [102]
    - Reduction [90]
    - more
  - Statistics [125]
  - Substances, Preparations & Products [277]
  - Biological Material & Organisms for Animal U
  - Method Specification [36]
  - Animal Species [40]
  - Product Properties & Effects [62]
  - Product Testing & Assessment [20]
  - 3Rs Methods in the Life Sciences [6]
  - Animal Experiment [61]
  - 3Rs Relevant [5]
    - In Vitro Experimental Design [20]
    - In Vivo Experimental Design [5]
    - Animal Condition, Physiological or Psycholog
    - Animal Care & Handling [3]
    - Toxic Actions of Substances [7]
    - Unclassified [390]
  - Find related categories ...
  - My last 5 queries
  - Clipboard [0]

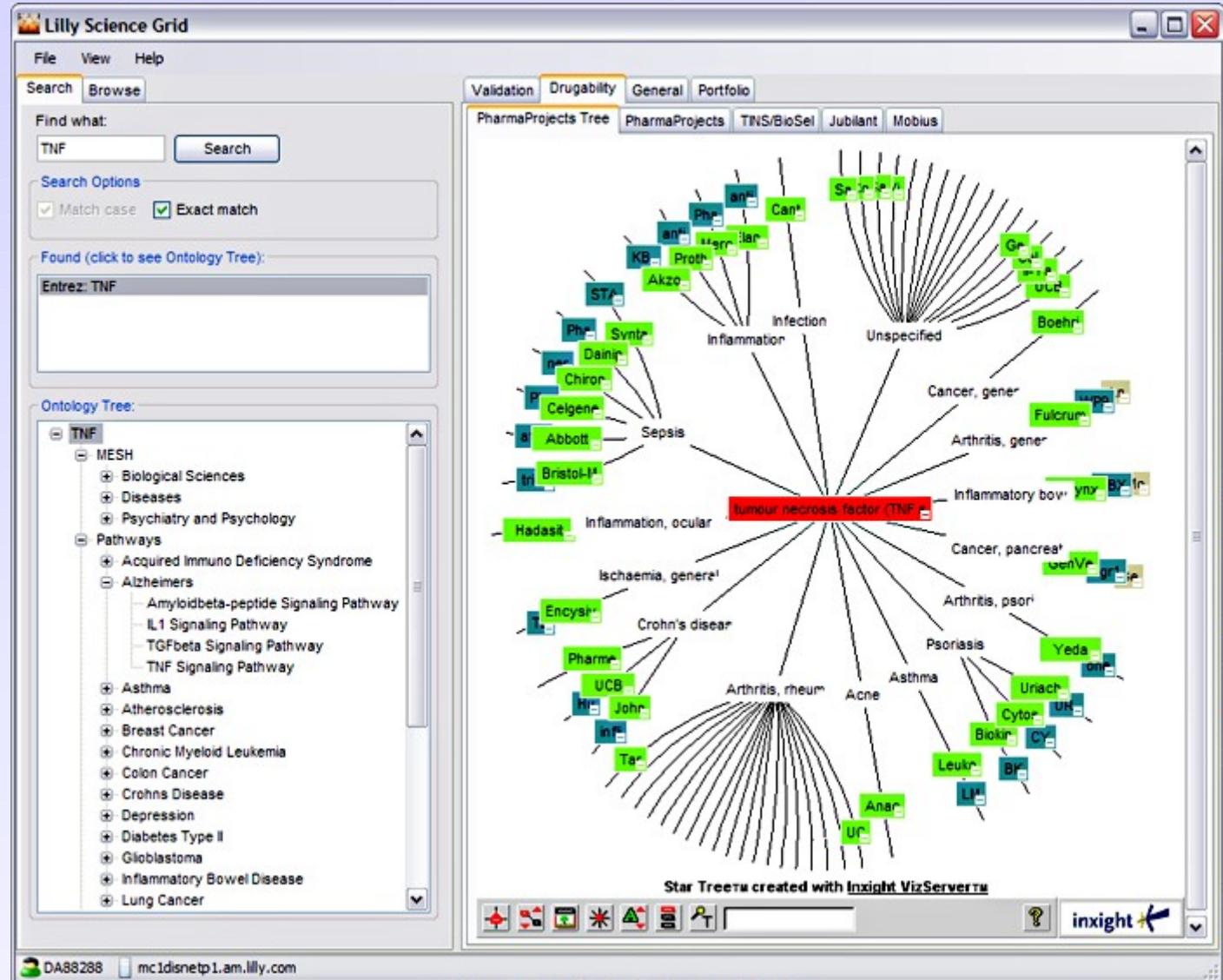
The main content area displays six search results:

- 2: Microvascular decompression of cochleovestibular nerve.**  
PMID: 18389269 Related Articles  
Yap L et al., Eur Arch Otorhinolaryngol, 2008  
This report provides a review of all the published studies on MVD of the eighth (8th) nerve in alleviating cochleovestibular symptoms and presents three additional patients who underwent MVD of the eighth nerve for **tinnitus** or vertigo.
- 3: Algorithm for evaluation of pulsatile tinnitus.**  
PMID: 18368578 Related Articles  
Mattox DE et al., Acta Otolaryngol, 128 (4): 427-31, 2008  
Among patients with arterial **tinnitus**, carotid atherosclerotic disease was the most common.
- 4: Functional imaging of unilateral tinnitus using fMRI.**  
PMID: 18368576 Related Articles  
Lanting CP et al., Acta Otolaryngol, 128 (4): 415-21, 2008  
This article shows that the inferior colliculus plays a key role in unilateral subjective **tinnitus**.
- 5: Pros and cons of tinnitus retraining therapy.**  
PMID: 18368566 Related Articles  
Hatanaka A et al., Acta Otolaryngol, 128 (4): 365-8, 2008  
A significant reduction in the **Tinnitus Handicap Inventory (THI)** was obtained as early as 1 month after implementation of **tinnitus** retraining therapy (TRT).
- 6: Mass casualty incident management triage injury distribution of casualties and**

# Example: Eli Lilly's target assessment tool

Prioritization of drug target, integrating data from different sources and formats

Integration, search via ontologies (proprietary and public)

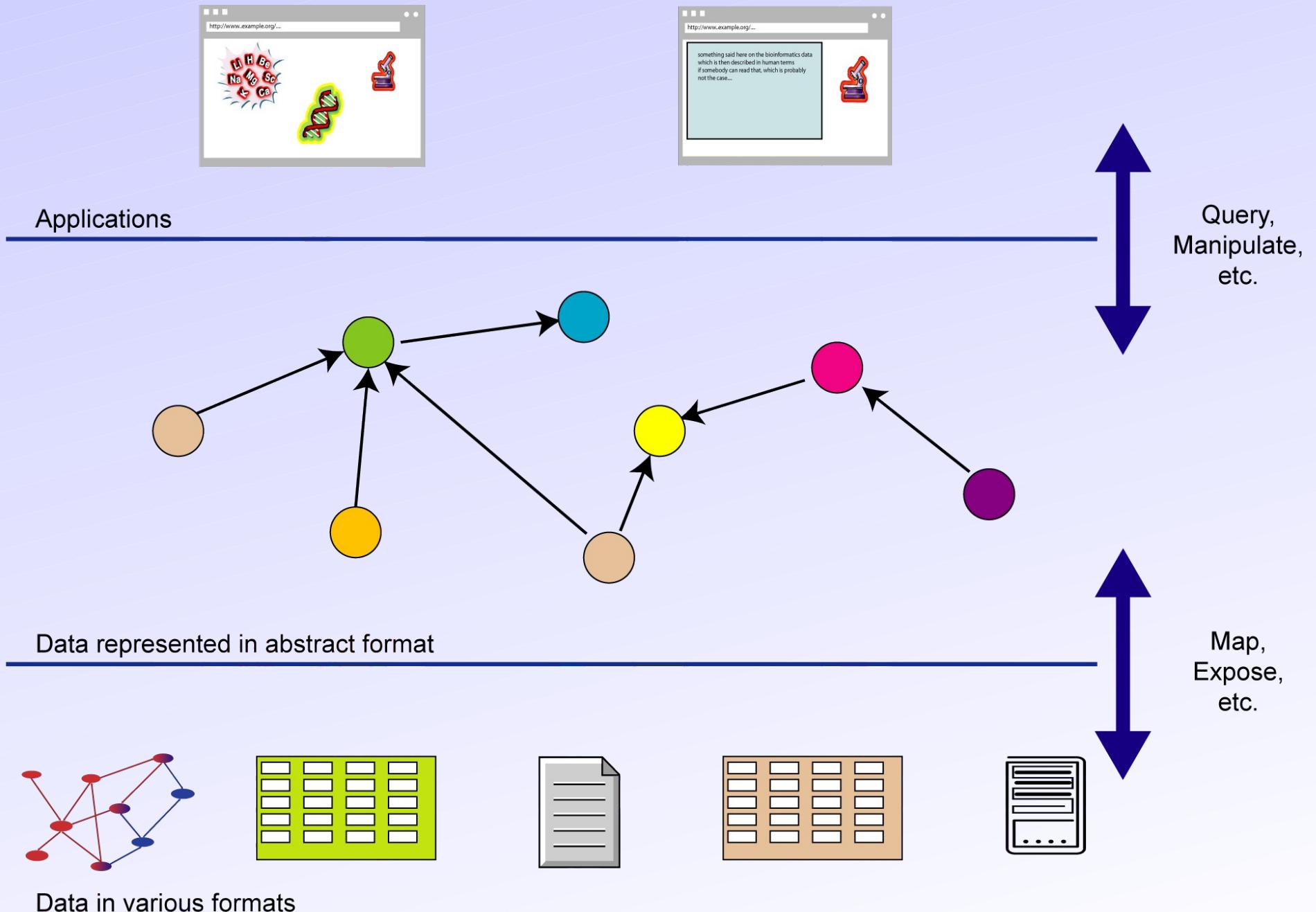


# ***What have we achieved? (putting all this together)***

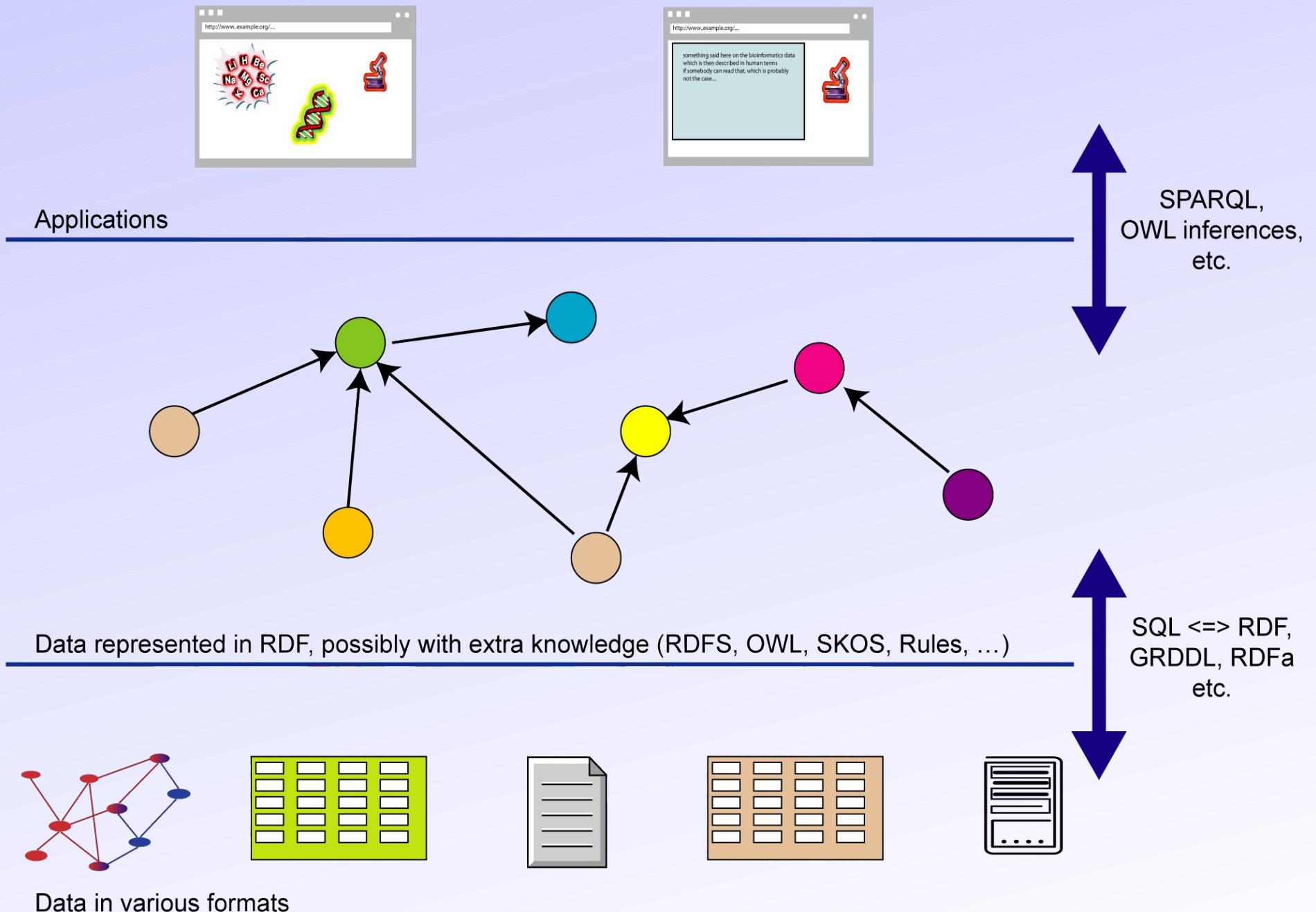
# *Other SW technologies*

- There are other technologies that we do not have time for here
  - find RDF data associated with general URI-s: POWDER
  - bridge to thesauri, glossaries, etc: SKOS
  - use Rule engines on RDF data

# Remember the integration example?



# Same with what we learned



# Example: personalized tourist itinerary

The screenshot shows a personalized tourist itinerary for Zaragoza on June 17, 2008. The itinerary is divided into Morning and Afternoon sessions. The Morning session includes visits to the Basilica of the Pilar, Ibercaja Camón Aznar Museum, Cathedral of San Salvador or La Seo, Caesaraugusta Forum Museum, Caesaraugusta River Port Museum, and Molins house. The Afternoon session includes visits to various churches and landmarks like the Church of la Mantería, Church of San Ildefonso or de Santiago el Mayor, Church of Santo Tomás de Aquino, Central market, Samartínana Fountain, Church of Santa Isabel de Portugal or San Cayetano, Church of San Felipe and Santiago el Menor, Church of San Juan de los Panetes, Church of Santa Cruz, and Church of Santa María Magdalena. A map of Zaragoza shows the locations of these sites with numbered pins. A detailed description of the Basilica of the Pilar is provided, mentioning its history and accessibility.

Integration of relevant data in Zaragoza (using RDF and ontologies)

Use rules on the RDF data to provide a proper itinerary

# *Available documents, resources*

# ***Available specifications: Primers, Guides***

- The “RDF Primer” and the “OWL Guide” give a formal introduction to RDF(S) and OWL
- GRDDL and RDFa Primers have also been published
- The W3C Semantic Web Activity Homepage has links to all the specifications:
  - <http://www.w3.org/2001/sw/>

# “Core” vocabularies

- There are also a number widely used “core vocabularies”
  - Dublin Core: about information resources, digital libraries, with extensions for rights, permissions, digital right management
  - FOAF: about people and their organizations
  - DOAP: on the descriptions of software projects
  - SIOC: Semantically-Interlinked Online Communities
  - vCard in RDF
  - ...
- One should never forget: ontologies/vocabularies must be shared and reused!

## *Some books*

- G. Antoniu and F. van Harmelen: Semantic Web Primer, 2<sup>nd</sup> edition in 2008
- D. Allemang and J. Hendler: Semantic Web for the Working Ontologist, 2008
- Jeffrey Pollock: Semantic Web for Dummies, 2009
- ...

See the separate Wiki page collecting book references:  
<http://esw.w3.org/topic/SwBooks>

## *Further information*

- Planet RDF aggregates a number of SW blogs:
  - <http://planetrdf.com/>
- Semantic Web Interest Group
  - a forum developers with archived (and public) mailing list, and a constant IRC presence on freenode.net#swig
  - anybody can sign up on the list:
    - <http://www.w3.org/2001/sw/interest/>

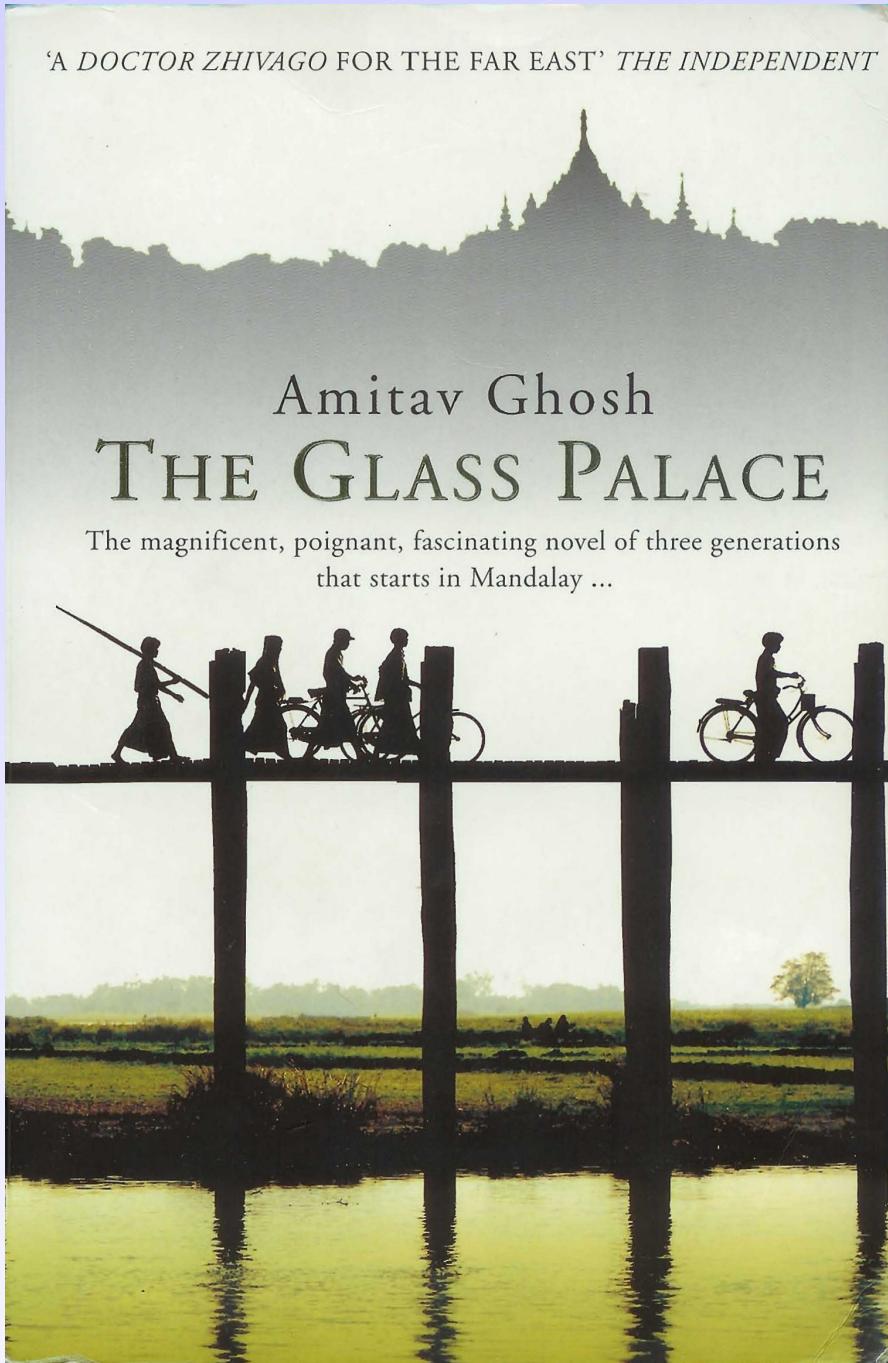
# **Lots of Tools (*not* an exhaustive list!)**

- Categories:
  - Triple Stores
  - Inference engines
  - Converters
  - Search engines
  - Middleware
  - CMS
  - Semantic Web browsers
  - Development environments
  - Semantic Wikis
  - ...
- Some names:
  - Jena, AllegroGraph, Mulgara, Sesame, flickurl, ...
  - TopBraid Suite, Virtuoso environment, Falcon, Drupal 7, Redland, Pellet, ...
  - Disco, Oracle 11g, RacerPro, IODT, Ontobroker, OWLIM, Tallis Platform, ...
  - RDF Gateway, RDFLib, Open Anzo, DartGrid, Zitgist, Ontotext, Protégé, ...
  - Thetus publisher, SemanticWorks, SWI-Prolog, RDFStore...
  - ...

# Conclusions

- The Semantic Web is about creating a Web of Data
- There is a great and very active user and developer community, with new applications
  - witness the size and diversity of this event

# By the way: the book is real 😊



# Thank you for your attention!

These slides are also available on the Web:

<http://www.w3.org/2009/Talks/0615-SanJose-tutorial-IH/>

