

Improving Search  
with  
Semantic Technologies:  
Current Research Directions

*Hugo Zaragoza*

**Yahoo! Research (Barcelona)**

# Yahoo! by numbers (April, 2007)

- **500 million users** of Yahoo! branded services (**1/2 users** online).
- Most visited site online
  - most visits (nearly 4 billion visits)
  - targets online time share: 12% in America and 8% worldwide.
  - #1 Web mail provider in the world (243 million users)
  - #1 online news destination
  - (leads all competitors in audience reach, frequency and engagement)
- Social media properties:
  - 115 million unique visitors worldwide.
  - Flickr: more than 450 million photo, 1 million uploaded daily.
  - Yahoo! Answers: 250 million answers worldwide.
  - Nearly 1 in 10 Internet users is a member of a Yahoo! Groups.

*(Flickr, delicious, Answers, 360, Video, MyBlogLog, Jumpcut and Bix)*

(comScore WorldMetrix Feb.07 & Yahoo! Internal Data )

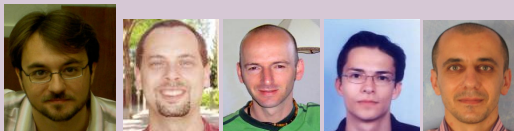
Hugo Zaragoza, ESTC 2008. (2)

# Yahoo! Research

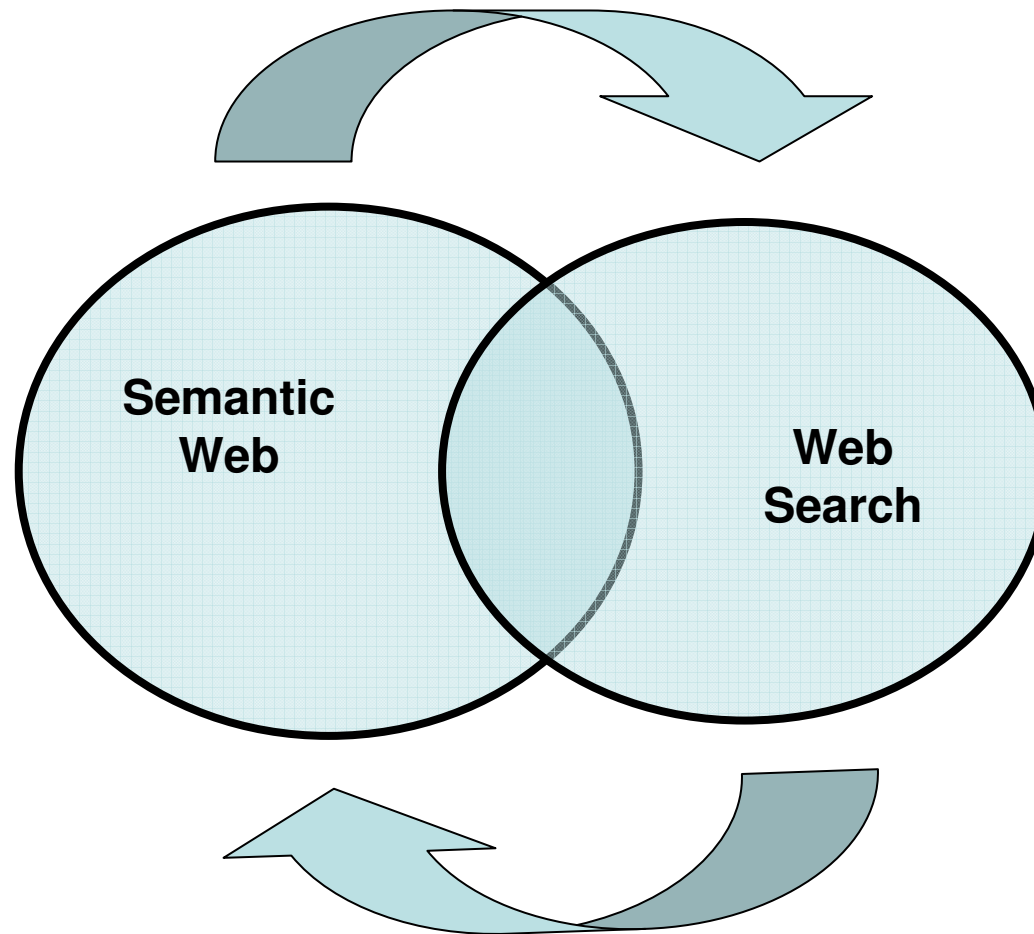


*The mission of Yahoo! Research is to develop the world-class science that will deliver the next generation of businesses to the company.*

## *Natural Language Retrieval group*



# Why should you care about search?



# Plan

- Search Monkey

*(today)*

- Web Search: where are we going?

*(tomorrow)*

- Current Research Directions

*(the future...)*

what is it?



SearchMonkey

an open platform for using structured data to build more useful and relevant search results

Before

Topics for **Getting Pregnant** - BabyCenter

Find out how to boost your chances of **getting pregnant**, what you can do if you're having a problem conceiving, and more. ... do before you try to **get pregnant** ...

[www.babycenter.com/getting-pregnant](http://www.babycenter.com/getting-pregnant) - 91k - [Cached](#)



After

Topics for **Getting Pregnant** - BabyCenter



[Tools](#)  
[Before you Try](#)  
[Active Trying](#)  
[Having Trouble](#)

BabyCenter's getting **pregnant** tools and information can boost your chances of conception by helping you chart your cycle, read your cervical mucus, and pinpoint ovulation.

[www.babycenter.com/getting-pregnant](http://www.babycenter.com/getting-pregnant) - 76k - [Cached](#)



WebMD **Allergies** Health Center - Find allergy information and latest ...

Information and articles on the diagnosis, symptoms, treatment, and the prevention of **allergies**.

[www.webmd.com/diseases\\_and\\_conditions/allergies.htm](http://www.webmd.com/diseases_and_conditions/allergies.htm) - 132k - [Cached](#)



WebMD **Allergies** Health Center - Find allergy information and latest ...



[Overview](#)  
[Symptoms](#)  
[Diagnosis](#)  
[Treatment](#)

**Allergies** are an abnormal response of the immune system. People who have allergies have an immune system that reacts to a usually harmless substance in the environment. This substance (pollen, mold, dander, etc.) is called an allergen.

[www.webmd.com/allergies](http://www.webmd.com/allergies) - 111k - [Cached](#)



**Italy Travel Guide** and Travel Information - Lonely Planet

Lonely Planet **Italy** includes information on events, attractions, activities, and transportation for the independent traveler.

[www.lonelyplanet.com/worldguide/destinations/europe/italy](http://www.lonelyplanet.com/worldguide/destinations/europe/italy) - 57k - [Cached](#)



**Italy Travel Guide: Overview** - Lonely Planet WorldGuide



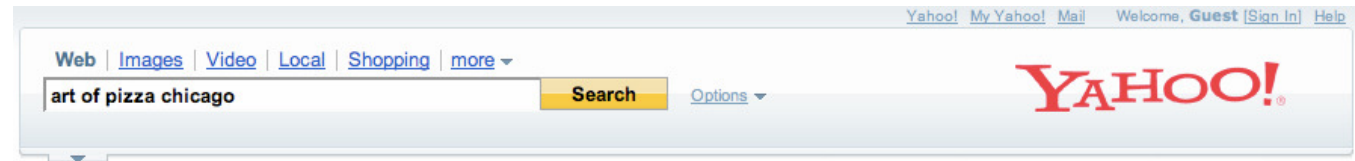
[Overview](#)  
[Money & Costs](#)  
[Itineraries](#)  
[Sights](#)

When to go: **Italy** is at its best in spring (April-May) and autumn (October-November). During these seasons, the scenery is beautiful, the temperatures are pleasant and are relatively few crowds.

[www.lonelyplanet.com/worldguide/destinations/europe/italy](http://www.lonelyplanet.com/worldguide/destinations/europe/italy) - 46k - [Cached](#)



# Enhanced Result: what might a developer build?



1 - 10 of about 8,140,000 for art of pizza chicago (About this page) - 0.23 sec.

deep  
links

image  
s

name/value  
pairs or  
abstract

[The Art of Pizza - Chicago, IL : Citysearch.com](#)  
Get details on The Art of Pizza - Chicago, IL, at Citysearch - over 1 million user reviews & edit/update about local businesses.  
[www.citysearch.com/profile/3735139](http://www.citysearch.com/profile/3735139)

[The Art of Pizza - Chicago, IL, 60657-3035 - Citysearch](#)  
... information, directions, and reviews on The Art of Pizza and other 5-Star Restaurants in Chicago ... The Art of Pizza. 3033 N Ashland Ave. Chicago, IL 60657 ...  
[chicago.citysearch.com/profile/.../chicago\\_il/the\\_art\\_of\\_pizza.html](http://chicago.citysearch.com/profile/.../chicago_il/the_art_of_pizza.html) - 76k - [Cached](#)

[The Art of Pizza - Lakeview - Chicago, IL 60657](#)



[Reviews](#)  
[Photos](#)  
[Send to a friend](#)  
[Send to Phone](#)

Ratings: ★★★★★ (173)  
Address: 3033 N Ashland Ave, Chicago, IL  
Phone: (773) 327-5600  
Price Range: \$

[www.acmereviews.com/biz/the-art-of-pizza-chicago](http://www.acmereviews.com/biz/the-art-of-pizza-chicago) - 145k - [Cached](#)

[The Art of Pizza | Metromix Chicago](#)

This modest storefront eatery's deep-dish pizza was voted Chicago's best by the Tribune. ... Rate The Art of Pizza "is not worth it" credit from logan ...  
[chicago.metromix.com/restaurants/italian/.../140369/content](http://chicago.metromix.com/restaurants/italian/.../140369/content) - [Cached](#)

[Art of Pizza: Chicago Reader Restaurants: Restaurant Raters' Comments](#)

Art of Pizza. 3033 N. Ashland Ave. Chicago ... Chicagoan, the Art of Pizza is one of the pizza greats in Chicago, and a relatively hidden gem. ...  
[www.chicagoreader.com/cgi-bin/rm/comments.cgi?numb=2736](http://www.chicagoreader.com/cgi-bin/rm/comments.cgi?numb=2736) - [Cached](#)

[Art of Pizza on Centerstage Chicago - Art of Pizza : 3033 N. Ashland ...](#)

Art of Pizza on Centerstage, Publisher of Honest Info By Chicago, For Chicago ... Eateries like Lakeview's Art of Pizza prove that a food lover's paradise can be ...  
[centerstagechicago.com/restaurants/art-of-pizza.html](http://centerstagechicago.com/restaurants/art-of-pizza.html)

## SPONSOR RESULTS

[Chicago Pizza Chicago](#)

Pizza Now - Call Today Chicago Style, World Class Taste.  
[dexonline.com/ledos-pizza](http://dexonline.com/ledos-pizza)

[Nick & Bruno's Pizzeria](#)

Chicago area pizza restaurant Home of the Monster Pizza.  
[Nick-Brunos.com](http://Nick-Brunos.com)

[Pizza Place in Chicago](#)

Find & review local restaurants and cafes. Free search.  
[Chicago.Citysearch.com](http://Chicago.Citysearch.com)

[Pizza Art](#)

Find Pizza Art and Compare prices at Smarter.com.  
[www.smarter.com](http://www.smarter.com)

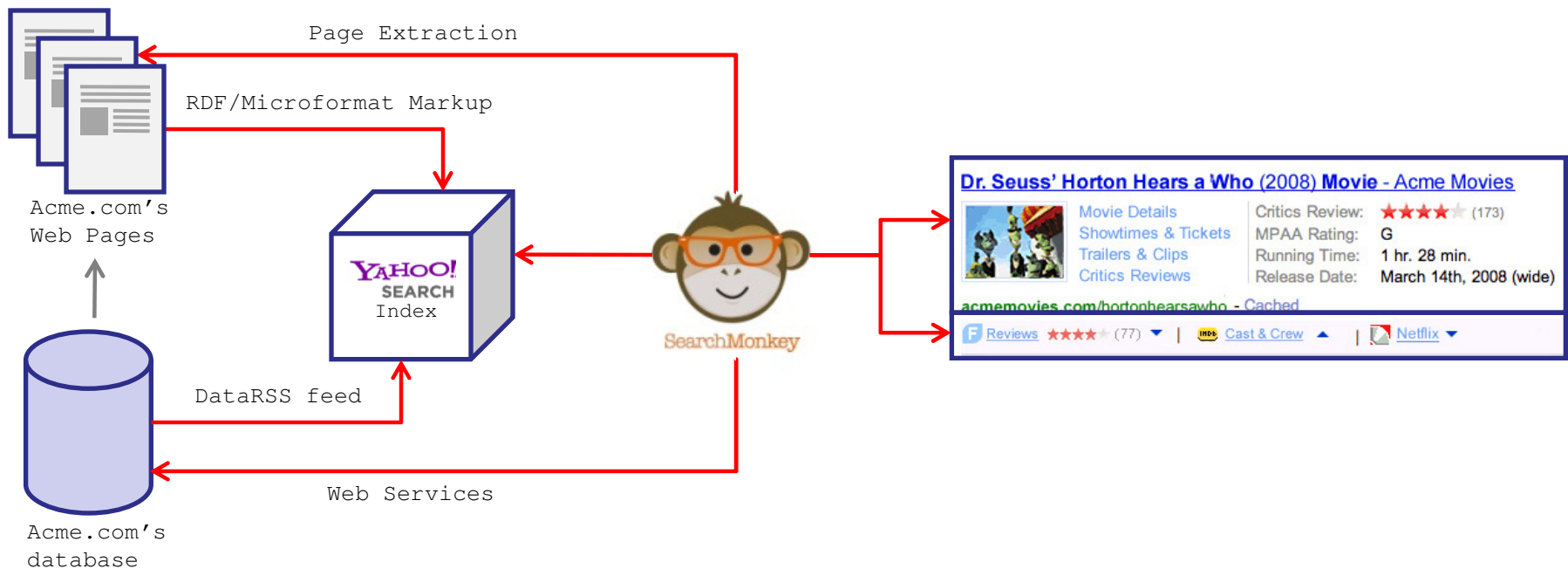
[See your message here...](#)



# How does it work?



- 1 site owners/publishers share structured data with Yahoo!.
- 2 site owners & third-party developers build SearchMonkey apps.
- 3 consumers customize their search experience with Enhanced Results or Infobars





# Infobar: what might a developer build?



Yahoo! My Yahoo! Mail Welcome, Guest [Sign In] Help

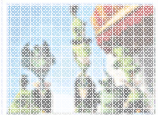
Web | Images | Video | Local | Shopping | more ▾

Dr. Seuss Horton Hears a Who Movie  Options ▾

1 - 10 of about 17,800,000 for Dr. Seuss' Horton Hears a Who Movie (About this page) - 0.33 sec.

**Dr. Seuss' Horton Hears A Who**  
[www.hortonmovie.com/splash.html](http://www.hortonmovie.com/splash.html) - [Cached](#)

**Dr. Seuss' Horton Hears a Who (2008) Movie - Acme Movies**

	<a href="#">Movie Details</a> <a href="#">Showtimes &amp; Tickets</a> <a href="#">Trailers &amp; Clips</a> <a href="#">Critics Reviews</a>	Critics Review: ★★★★★ (173) MPAA Rating: G Running Time: 1 hr. 28 min. Release Date: March 14th, 2008 (wide)
---	---	---

[acmemovies.com/hortonhearsawho](http://acmemovies.com/hortonhearsawho) - [Cached](#)

[Reviews](#) ★★★★★ (77) | [Cast & Crew](#) | [Netflix](#)

**Movie Reviews - Flixster**

 ★ **Top Critic** Seuss lovers may now heave a sigh of relief. Unlike those behind the recent live-action grotesqueries *The Grinch* and *Cat in the Hat*, the makers of the *Horton Hears a Who!* do well by the great, good metric of metrical writing.

Moviebuff71  
San Francisco, CA [See 76 more reviews](#)

**Dr. Seuss' Horton Hears a Who! Movie Times**  
Find Dr. Seuss' Horton Hears a Who! movie times, buy tickets, and get a map to the theater at Fandango.com. ... Dr. Seuss' Horton Hears a Who!Times and ...  
[www.fandango.com/dr.seusshortonhearsawho\\_102891/movietimes](http://www.fandango.com/dr.seusshortonhearsawho_102891/movietimes) - [Cached](#)

**Dr. Seuss' Horton Hears a Who! Movie Reviews, Pictures - Rotten Tomatoes**  
Dr. Seuss' Horton Hears a Who! movie reviews, trailers - Check out Rotten ... MOVIES / IN

SPONSOR RESULTS  
[Hottest Movies of 2008](#)  
Access to All the Info You Need On Horton Hears a Who.  
[www.RottenTomatoes.com](http://www.RottenTomatoes.com)

[See your message here...](#)

Pull in data from any web service



# Infobar: what might a developer build?



Yahoo! My Yahoo! Mail Welcome, Guest [Sign In] Help

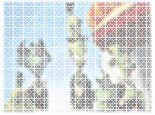
Web | Images | Video | Local | Shopping | more ▾

Dr. Seuss Horton Hears a Who Movie  Options ▾

1 - 10 of about 17,800,000 for Dr. Seuss' Horton Hears a Who Movie (About this page) - 0.33 sec.

[Dr. Seuss' Horton Hears A Who](#)  
www.hortonmovie.com/splash.html - [Cached](#)

[Dr. Seuss' Horton Hears a Who \(2008\) Movie - Acme Movies](#)

 [Movie Details](#) | [Showtimes & Tickets](#) | [Trailers & Clips](#) | [Critics Reviews](#)

Critics Review: ★★★★★ (173)  
MPAA Rating: G  
Running Time: 1 hr. 28 min.  
Release Date: March 14th, 2008 (wide)

acmemovies.com/hortonhearsawho - [Cached](#)

[Reviews](#) ★★★★★ (77) | [Cast & Crew](#) | [Netflix](#)

**Cast & Crew: Horton Hears a Who! (2008) - Internet Movie Database**

Actor	Role
<a href="#">Jim Carrey</a>	Horton (voice)
<a href="#">Steve Carell</a>	The Mayor of Whoville (voice)
<a href="#">Carol Burnett</a>	Kangaroo (voice)
<a href="#">Dan Fogler</a>	Councilman / Yummo Wickersham
<a href="#">Isla Fisher</a>	Dr. Mary Lou Larue (voice)

[Dr. Seuss' Horton Hears a Who! Movie Times](#)  
Find Dr. Seuss' Horton Hears a Who! movie times, buy tickets, and get a man to the theater at

SPONSOR RESULTS  
[Hottest Movies of 2008](#)  
Access to All the Info You Need On  
orton Hears a Who.  
www.RottenTomatoes.com

[See your message here...](#)

Open  
canvas for  
HTML



# Infobar: what might a developer build?



Yahoo! My Yahoo! Mail Welcome, Guest (Sign In) Help

Web | Images | Video | Local | Shopping | more ▾

Dr. Seuss Horton Hears a Who Movie  Options ▾

1 - 10 of about 17,800,000 for Dr. Seuss' Horton Hears a Who Movie (About this page) - 0.33 sec.

[Dr. Seuss' Horton Hears A Who](#)  
www.hortonmovie.com/splash.html - [Cached](#)

[Dr. Seuss' Horton Hears a Who \(2008\) Movie - Acme Movies](#)

	Movie Details	Critics Review: ★★★★★ (173)
	Showtimes & Tickets	MPAA Rating: G
	Trailers & Clips	Running Time: 1 hr. 28 min.
	Critics Reviews	Release Date: March 14th, 2008 (wide)

acmemovies.com/hortonhearsawho - [Cached](#)

F Reviews ★★★★★ (77) | IMDb Cast & Crew |  Netflix ▲

**Horton Hears a Who! (2008) - Netflix** **NETFLIX**

**Dr Seuss's** beloved story about an elephant that discovers a tiny society existing on a spot of dust springs to life in this vibrant animated adaptation from Academy Award-winning animator Chuck Jones. Agreeing to protect **Who**-ville from harm, **Horton** the elephant inspires the **Whos** to make their presence known among all the inhabitants of the jungle and champion the idea that "a person is a person, no matter how small."

Enjoyed By Members Who Enjoyed

- [The Muppet Show: Season 2 \(4-Disc Series\)](#)
- [101 Dalmatians \(Animated\)](#)
- [It's the Great Pumpkin, Charlie Brown](#)





## Everything SearchMonkey:

[developer.yahoo.com/searchmonkey](http://developer.yahoo.com/searchmonkey)

### mailing lists

- [searchmonkey-developers@yahoogroups.com](mailto:searchmonkey-developers@yahoogroups.com)
- [searchmonkey-siteowners@yahoogroups.com](mailto:searchmonkey-siteowners@yahoogroups.com)

### forums

- <http://suggestions.yahoo.com/searchmonkey>
- <http://suggestions.yahoo.com/datarss>

### data/partners



...and all sites with supported microformats, or eRDF markup!



# Plan

- Search Monkey *(today)*
- Web Search: where are we going? *(tomorrow)*
- Current Research Directions *(the future...)*

Web Images Video Local Shopping more  
**YAHOO! SEARCH** papa johns Search

Answers

Search Results

1 - 10 of about 761,000 for papa johns - 0.09 sec. (About this page)

Did you mean: [papa john's](#)

- Papa John's** (Nasdaq: [PIZZA](#))  
Official site for the pizza delivery and carry-out chain **Papa John's**. Includes store locator, nutritional info, and franchise and employment opportunities.  
[Papa John's near you](#)  
[www.papajohns.com](#) - 14k - [Cached](#) - [More from this site](#)
- [Welcome to Papa John's Online Ordering](#)  
Order **Papa Johns** Pizza online today. It's easy to register and order online. ... Welcome to **Papa John's** Online Ordering. New online customer? Click Here to get ...  
[www.papajohnsonline.com](#) - 9k - [Cached](#) - [More from this site](#)
- [Papa John's Menu](#)  
Get a glimpse of **Papa** John's basic menu for both our restaurants and online ordering. ... **Papa** John's does not add monosodium glutamate to any of our pizza toppings. ...  
[www.papajohns.com/menu/index.htm](#) - 49k - [Cached](#) - [More from this site](#)
- [Papa John's Restaurant Locator](#)  
Find the neighborhood **Papa Johns** location in your area. You can order online or by phone. ...  
©Copyright 2007 **Papa** John's International Inc. All Rights Reserved ...  
[www.papajohnsonline.com/restlocator.jsp](#) - 18k - [Cached](#) - [More from this site](#)
- [Papa John's Pizza UK - America's Number 1](#)  
Better Ingredients, Better Pizza - UK Delivery Operator Of The Year 2004 And 2005 ... Privacy Statement | Terms and Condition | **Papa** John's (GB) Ltd 2006 ...  
[www.papajohns.co.uk](#) - 10k - [Cached](#) - [More from this site](#)

SPONSOR RESULTS

[Papa John's Official Site](#)  
\$10.40 Tax Special - Large Pizza  
With 5 Toppings. Order Online.  
[www.papajohns.com](#)

[Papa Johns For Less](#)  
Looking for **papa johns**? Find exactly what you want today.  
[www.eBay.com](#)

[Free Pizza](#)  
Get Free Pizza for a Year. Hurry Offer Ends Soon. Free Pizza.  
[www.totally-free-stuff.org](#)

[See your message here...](#)



view larger map

Category:

- Pizza (5)
Restaurants (5)
Carry Out & Take Out (4)
Catering Services (1)
Food Delivery Services (1)
Food Services (1)
Italian Restaurants (1)

Rating:

- (rating key)
5 Stars (1)
4 Stars and above (2)
3 Stars and above (3)
2 Stars and above (3)
1 Star and above (4)

Name:

(650) 342-9392, San Mateo, CA
www.amicis.com Map | Directions | Send to Phone

- Pizz'a Chicago Contact Us Today for Additional Information
(408) 283-9400, 155 W San Fernando St, San Jose, CA
www.pizzachicago.com Map | Directions | Send to Phone

Sorted by: Top Results | Distance | Name | Rating | User Recommended

Printable Version

1. Papa John's Pizza

(408) 370-7272 1824 Saratoga Ave, Saratoga, CA 1.62 mi
Map | Directions | Send to Phone | Save to Collection

...near the Burrito Factory about 3 blocks away, their pizza is much better.

Papa... more

See all: Carry Out & Take Out - Pizza - Restaurants

pizzabizinc.com/



2. Papa John's

(408) 879-0202 1920 Camden Ave, San Jose, CA 4.72 mi
Map | Directions | Send to Phone | Save to Collection

...Papa Johns is a take out and delivery restaurant with a red and white counter.

They... more

See all: Pizza - Restaurants

www.papajohns.com/



3. Papa John's

(408) 615-7272 1998 Homestead Rd, Santa Clara, CA 5.34 mi
Map | Directions | Send to Phone | Save to Collection

...Papa Johns is a take out and delivery restaurant with a red and white counter.

They... more





Web Images Video News Maps more »

sfo bos apr 25

Search

Advanced Search Preferences

Web Results 1 - 10 of about 652,000 for sfo bos apr 25. (0.10 sec)

Flights from San Francisco, CA to Boston, MA

Departing: 04/25 Returning: 05/02 CheapTickets - Expedia - Hotwire - Orbitz - Priceline - Travelocity

FoRK Archive: Re: Oh, the trials of LaMacchia

UA 04-APR-97 0033 BOS SFO 2704 2704 5408 UA 01-APR-97 0160 LAX BOS 2611 2611 5222 UA 28-MAR-97 0159 BOS LAX 2611 2611 5222 UA 25-MAR-97 0160 LAX BOS 2611 ... www.xent.com/FoRK-archive/may97/0789.html - 10k - Cached - Similar pages - Note this

San Francisco (SFO) to Boston (BOS) departing Apr 2007

San Francisco (SFO) to Boston (BOS) departing Apr 2007 Get this by email Get this ... Top 25 Airport Idea. add another. World Cities, United States, Europe ... www.pinpointtravel.com/h/buzz/view?dt=A&code=SFO&dest=BOS&tm=200704 - 66k - Cached - Similar pages - Note this

ESPN.com - MLB - Box Score - Twins at Royals

MLB Scoreboard - April 25, 2006. - American League - BOS, 8, F ... Batters faced: S Elarton 25; L Hudson 3; A Sisco 3; A Burgos 4 ... sports.espn.go.com/mlb/boxscore?gameId=260425107 - 130k - Cached - Similar pages - Note this

ESPN.com - MLB - Box Score - Devil Rays at Yankees

MLB Scoreboard - April 25, 2006. - American League - BOS, 8, F. CLE, 6. TAM, 1. NYY, 9, F. BAL, 3 ... STL, 6, F. ARI, 7, F. SDG, 0. NYM, 4, F. SFO, 1 ... sports.espn.go.com/mlb/boxscore?gameId=260425110 - 130k



# Beyond traditional “Web Search”

- We move from a web of pages to a web of objects
- Objects are people, places, business, flights... and their relationships.
- User Intents are satisfied by juxtapositions of objects

# Beyond traditional “Web Search”

- Search is no longer about finding documents
  - It is an interface for web-mediated goals
- Old battles
  - Precision of navigational queries (solved)
  - Document crawling and spam (!)
  - Indexing and retrieval (!)
  - Result relevance (!)

# Hard problems to be solved

- How do you model intent?
  - What is the right abstraction?
  - What is the right granularity?
  - What are the top intents?
- How do you measure user happiness?
  - Transaction?
  - User studies? Interviews?

# Hard problems to be solved

- What is the right data model for the web of objects?
- What is the automated framework for relevance?
- What are the ranking models that can attain it?


# Plan

- Search Monkey *(today)*
- Web Search: where are we going? *(tomorrow)*
- Current Research Directions *(the future...)*
  - *MicroSearch*
  - *Learning tags*
  - *Searching Objects*

# Micro Search, example: ivan herman

query: ivan herman

**Search Results**

- [Ivan's blog](http://www.ivan-herman.net/)  
http://www.ivan-herman.net/ - 54k Update metadata
- [Ivan Herman](http://www.w3.org/People/Ivan/)  
Ivan Herman gives a talk on behalf of the China Office entitled "What is the ... Ben Adida, Elias Torres, and Ivan Herman give a tutorial entitled "RDFa: ...  
http://www.w3.org/People/Ivan/ - 17k Update metadata
- [Ivan Herman's Photos](http://www.ivan-herman.net/Photos/)  
  
name: Herman Iván, Ivan Herman  
homepage: <http://www.ivan-herman.net>  
My Photos. I like traveling and making photos. ... **Iván Herman**, 2007-12-16, [ivan@ivan-herman.net](mailto:ivan@ivan-herman.net) The images are licensed under a ...  
<http://www.ivan-herman.net/Photos/> - 4k Update metadata  
**is primary topic of**  
[Ivan's blog](http://www.ivan-herman.net/)  
<http://www.ivan-herman.net/> - 54k Update metadata
- [Ivan Herman](http://homepages.cwi.nl/~ivan/)  
Ivan Herman. Centrum voor Wiskunde en Informatica (CWI) Kruislaan 413, P.O. Box 94079, ...  
Email: [Ivan.Herman@cwi.nl](mailto:Ivan.Herman@cwi.nl) or [ivan@w3.org](mailto:ivan@w3.org). GnuPG: Public GnuPG key ...  
<http://homepages.cwi.nl/~ivan/> - 4k Update metadata
- [Ivan Herman - Technical Presentations](http://www.techpresentations.org/Ivan_Herman)  
Ivan Herman. From Technical Presentations. Jump to: navigation, search ... about **Ivan Herman** RDF feed. Homepage. <http://www.w3.org/People/Ivan/> + Works at ...  
[http://www.techpresentations.org/Ivan\\_Herman](http://www.techpresentations.org/Ivan_Herman) - 13k Update metadata
- [Ivan herman - ResearchIndex document query](http://citeseer.ist.psu.edu/cis?q=Ivan+Herman)  
ivan herman - scientific articles matching the query: **ivan herman** ... **Ivan Herman** graduated as applied mathematician in 1979 in ...  
<http://citeseer.ist.psu.edu/cis?q=Ivan+Herman> - 19k Update metadata
- [Flickr: Ivan Herman](http://www.flickr.com/people/ivan_herman/)  
formatted name: Ivan Herman  
address: Amsterdam (well... in a suburb), The Netherlands  
url: <http://www.ivan-herman.net>  
Flickr is almost certainly the best online photo management and sharing ... Testimonials. **Ivan Herman** doesn't have any testimonials yet. You ...  
[http://www.flickr.com/people/ivan\\_herman/](http://www.flickr.com/people/ivan_herman/) - 20k Update metadata

**Geolocation**

**Related pages based on metadata**

**Conferences he plans to attend and other events from homepage plus bio events from LinkedIn**



Peter Mika. <http://www.yr-bcn.es/demos/microsearch>

# MicroSearch, example: san francisco conference

The screenshot shows a Yahoo! search results page for the query "san francisco conference". The search bar at the top contains the text "san francisco conference" and the "Search" button is visible. Below the search bar, the "Search Results" section displays a list of nine search results. The first result is from Encyclopaedia Britannica, titled "San Francisco Conference -- Encyclopaedia Britannica". The second result is from Oracle OpenWorld, titled "San Francisco". The third result is from ad.tech, titled "ad.tech Interactive Media Conference". The fourth result is from Only In San Francisco, titled "The Official Visitors Site for San Francisco". The fifth result is from Gilbane, titled "San Francisco 2008". The sixth result is from South San Francisco Conference Center, titled "South San Francisco Conference Center". The seventh result is from American Heart Association, titled "5th Annual Conference on Arteriosclerosis, Thrombosis and Vascular Biology". The eighth result is from ISPI, titled "ISPI Annual Conference 2007, San Francisco, California". The ninth result is from Wikipedia, titled "Treaty of San Francisco - Wikipedia, the free encyclopedia".

On the right side of the search results, there is a map of the world with a "Sat" button selected. Below the map is a calendar view for the year 2007, showing events for various months. The calendar view includes a "Timeline © SIMILE" and shows events for May, June, July, August, and September. The events listed include "@media 2007 America", "Workshop: Transcending CSS", "Open Source Convention", "@media 2007 Europe", "Future of Web Design Europe", and "Future of Web Design".

At the bottom of the page, there is a search bar with the text "Find: native" and a "Done" button. The "Proxy: None" option is also visible.

Conferences in San Francisco date

# MicroSearch, example: greater st. peter

The image shows a screenshot of a Yahoo! search results page for the query "greater st. peter". The page displays a list of search results, with the first result highlighted. A map of the United States is visible in the background, showing the location of the search results. The search results are as follows:

- 1. [Greater St. Peter @ River's Edge](#)  
**Name:** Senior Pastor: Isaiah J. Waddy, M. DIV.  
**Address:** 9540 Fayetteville Road, Jonesboro,  
tel: +1-770-210-4100  
Add to your AddressBook  
Greater St. Peter African Methodist Episcopal Church in Jonesboro, Georgia: A Church for ...  
mercy of God, **Greater St. Peter** has made significant strides ...  
<http://www.stpeterame.org/index.html> - 8k - [Cached](#)
- 2. [Greater St. Peter Missionary Baptist Church](#)  
a good church, a church you can learn the word of God, Rev. Jesse King is the Pastor ...  
Greater St. Peter Baptist Church "The Hands of God The Church Ministries" ...  
<http://www.peterchurchministry.com/> - 32k - [Cached](#)
- 3. [Greater Saint Peter](#)  
Christian congregation  
<http://www.poweredby...>
- 4. [St Peter's RC Sch](#)  
... reach for **greater** dr  
Buffalo. Campus Outfit  
<http://www.stpeterc.or>
- 5. [Jobs at The YMCA](#)  
Jobs at The YMCA of G  
engine. ... ymca of gre  
<http://www.careerjet.co>
- 6. [St. Peter's Episcop](#)  
The Homepage of St. P  
community; to work to  
<http://www.stpeterschic>
- 7. [St. Louis Honda de](#)  
St. Louis Honda dealer  
new Honda & used Hon  
<http://www.stpetershor>
- 8. [Greater St. Peter E](#)  
Get details on Greater  
user reviews & editorial  
<http://www.citysearch.c>

The search results are displayed on a page with the following elements:

- Search bar: "greater st.peter"
- Search button: "Search"
- Advanced Search link: "Advanced Search"
- Search Results section: "Search Results"
- Map: "Map", "Hyb", "Sat"
- Callouts: "Call on Skype", "Add to your addressbook", "(other actions)"



# MicroSearch, towards intent: paris hilton

paris hilton - Yahoo! Search Results - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://search.yahoo.com/search?p=paris+hilton&ei=UTF-8&fr=moz2 paris hilton

Yahoo! My Yahoo! Mail Welcome, serendipity588 [Sign Out] Help

Web Images Video Local Shopping more

paris hilton Search Options

1 - 10 of about 191,000,000 for paris hilton (About this page) - 0.05 sec.

Also try: [paris hilton photos](#), [free hot paris hilton photos](#), [More...](#)

**Paris Hilton** - Image Results

More [paris hilton images](#)

Yahoo! Shortcut About

**Paris Hilton** - Wikipedia

User-edited article on **Paris Hilton**, the **Hilton** Hotel heiress, actress, and model. Includes biography, filmography, and discography.  
[en.wikipedia.org/wiki/Paris\\_Hilton](http://en.wikipedia.org/wiki/Paris_Hilton) - 121k - [Cached](#)

**Paris Hilton** - MySpace

Official MySpace page for **Paris Hilton** includes songs from her album, videos, and a blog.  
[www.myspace.com/parishilton](http://www.myspace.com/parishilton) - 159k - [Cached](#)

**Paris Hilton Zone**

**Paris Hilton** fansite offers news updates and photo galleries of the heiress and socialite, including magazine spreads and awards show appearances.  
[www.parishiltonzone.com](http://www.parishiltonzone.com) - 48k - [Cached](#)

Done Proxy: None

# MicroSearch, towards intent: hugo zaragoza

**YAHOO! RESEARCH** **micro SEARCH** Hugo Zaragoza Search Note: response times may vary between 2-15 seconds.

Search Results Persons: 69 Events: 52 Unfinished: 9 | 1 - 10 of about 100905 for Hugo Zaragoza - 13 sec. (About this page)

- Hugo Zaragoza | Yahoo! Research**  
Hugo Zaragoza is a researcher working on Information Retrieval at Yahoo! ... optimization Robertson, Stephen ; Zaragoza, Hugo, Information Retrieval, 2007 ...  
<http://research.yahoo.com/~hugo2> - 42k Update metadata
- Hugo Zaragoza**  
Machine Learning, Information Retrieval, Information Extraction, Language Models, Statistical Natural Language Processing ... Hugo Zaragoza has now left Microsoft ...  
<http://research.microsoft.com/users/cambridge/hugo/> - 5k Update metadata
- LinkedIn: Hugo Zaragoza**  
formatted name: Hugo Zaragoza  
address: Barcelona Area, Spain  
url: <http://grupoweb.upf.es/hugo/>  
Hugo Zaragoza's professional ... Hugo Zaragoza. Senior Researcher at Yahoo! Research (Barcelona) ... Hugo Zaragoza's Education. Universit  Pierre et ...  
<http://www.linkedin.com/in/hugozaragoza> - 14k Update metadata
- hugo zaragoza - ResearchIndex document query**  
hugo zaragoza - scientific articles matching the query: hugo zaragoza ... time I spent working with him and Hugo Zaragoza during an internship at MS Research, ...  
<http://citeseer.ist.psu.edu/cis?q=Hugo+Zaragoza> - 16k Update metadata
- HugoZaragozaArtFutura.wav - FileFront.com**  
HugoZaragozaArtFutura.wav from FileFront.com - Download HugoZaragozaArtFutura.wav ... Act of War Age of Empires Aliens vs Predator America's Army Armed Assault ...  
<http://files.filefront.com/HugoZaragozaArtFuturawav;8902507;/fileinfo.html> - 52k Update metadata
- DBLP: Hugo Zaragoza**  
Hugo Zaragoza, Henning Rode, Peter Mika, Jordi Atserias, Massimiliano Ciaranita, ... Michael J. Taylor, Hugo Zaragoza, Nick Craswell, Stephen Robertson, Chris Burges: ...  
<http://www.informatik.uni-trier.de/~ley/db/indices/a-tree/z/Zaragoza:Hugo.html> - 16k Update metadata

Map of Barcelona showing two markers for Hugo Zaragoza. Timeline below the map shows activity from 2005 to 2007.

# Plan

- Search Monkey *(today)*
- Web Search: where are we going? *(tomorrow)*
- Current Research Directions *(the future...)*
  - *MicroSearch*
  - *Learning to tag, tagging to learn.*
  - *Ranking Objects*

# Example

- About Wikipedia
- Community portal
- Recent changes
- Contact Wikipedia
- Donate to Wikipedia
- Help

search

toolbox

- What links here
- Related changes
- Upload file
- Special pages
- Printable version
- Permanent link
- Cite this article

languages

- Afrikaans
- العربية
- বাংলা
- Bân-lâm-gú
- Bosanski
- Brezhoneg
- Български
- Català
- Česky
- Cymraeg
- Dansk

**Pablo Ruiz Picasso** (October 25, 1881 – April 8, 1973), often referred to simply as **Picasso**, was a [Spanish painter](#) and [sculptor](#). His full name is **Pablo Diego José Francisco de Paula Juan Nepomuceno María de los Remedios Cipriano de la Santísima Trinidad Clito Ruiz y Picasso**.<sup>[1]</sup> One of the most recognized figures in 20th century [art](#), he is best known as the co-founder, along with [Georges Braque](#), of [cubism](#).

**Contents** [\[show\]](#)

## Biography

[\[edit\]](#)

Pablo Picasso was born in [Málaga, Spain](#), the first child of [José Ruiz y Blasco](#) and [María Picasso y López](#). He was christened with the names Pablo, Diego, José, Francisco de Paula, Juan Nepomuceno, María de los Remedios, and Cipriano de la Santísima Trinidad.<sup>[2]</sup> Picasso's father was a painter whose specialty was the naturalistic depiction of birds and who for most of his life was also a [professor](#) of art at the School of Crafts and a [curator](#) of a local museum. The young Picasso showed a passion and a skill for drawing from an early age; according to his mother,<sup>[3]</sup> his first word was "piz," a shortening of *lápiz*, the Spanish word for [pencil](#).<sup>[4]</sup> It was from his father that Picasso had his first formal academic art training, such as figure drawing and painting in oil. Although Picasso attended art schools throughout his childhood, often those where his father taught, he never finished his college-level course of study at the Academy of Arts

**Pablo Picasso**



Picasso (January 1962)

**Birth name** Pablo Diego José Francisco de Paula Juan Nepomuceno María de los Remedios Cipriano de la Santísima Trinidad Martyr Patricio Clito Ruiz y Picasso

**Born** [October 25, 1881](#)

 [Málaga, Spain](#)

**Died** [April 8, 1973](#) (aged 91)

 [Mougins, France](#)

# Statistical Taggers @ Y!Research Barcelona

- Name Entity Recogniser



- SuperSense Tagger.

- Massimiliano Ciaramita*

- Dependency Parsing



- 2nd best at CONL 2007

- Giuseppe Attardi (U. Pisa)*

- Semantic Role Labeling



- 2nd best at CONL 2008

- Mihai Surdeanu.*

Fast, realistic, agnostic

# Extending annotations

Pablo Picasso was born in Málaga, Spain.

PER

LOC

LOC

E:PERSON

GPE:CITY

GPE:COUNTRY

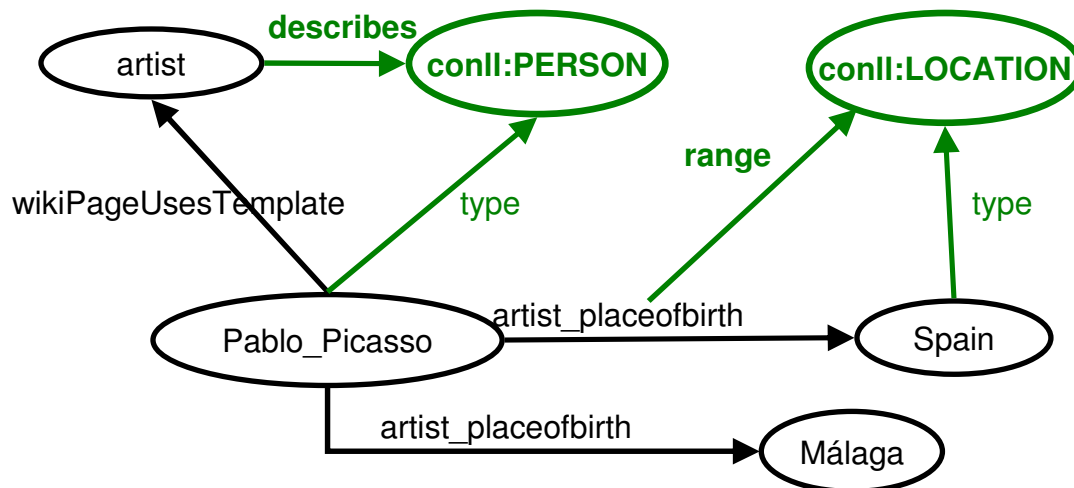
artist:name

artist:placeofbirth

artist:placeofbirth



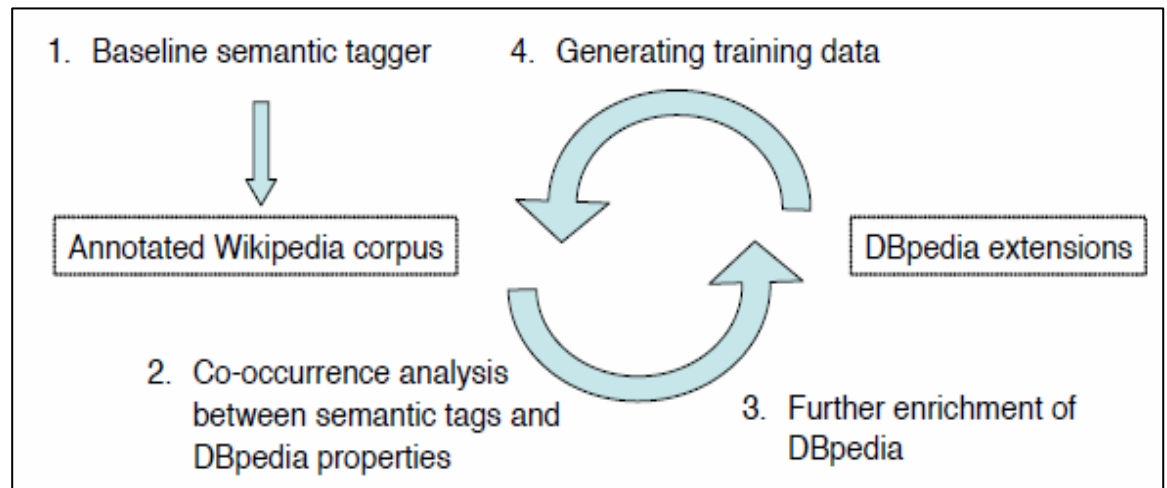
If most artists are persons, then let's assume all artists are persons.  
If most places of birth are locations, then let's assume all are.



[Mika, Zaragoza, Ciaramita, Atserias. IEEE AI 2008]

# Adaptation and improved tagging

- Generate training data based on what is learned
  - i.e. if Paris is also a place of birth, it is also a place
- Dealing with sparse data  
Paris is the capital city of France.
- Improvement
  - On news: +6.1%
  - On Wikipedia: +5.5%

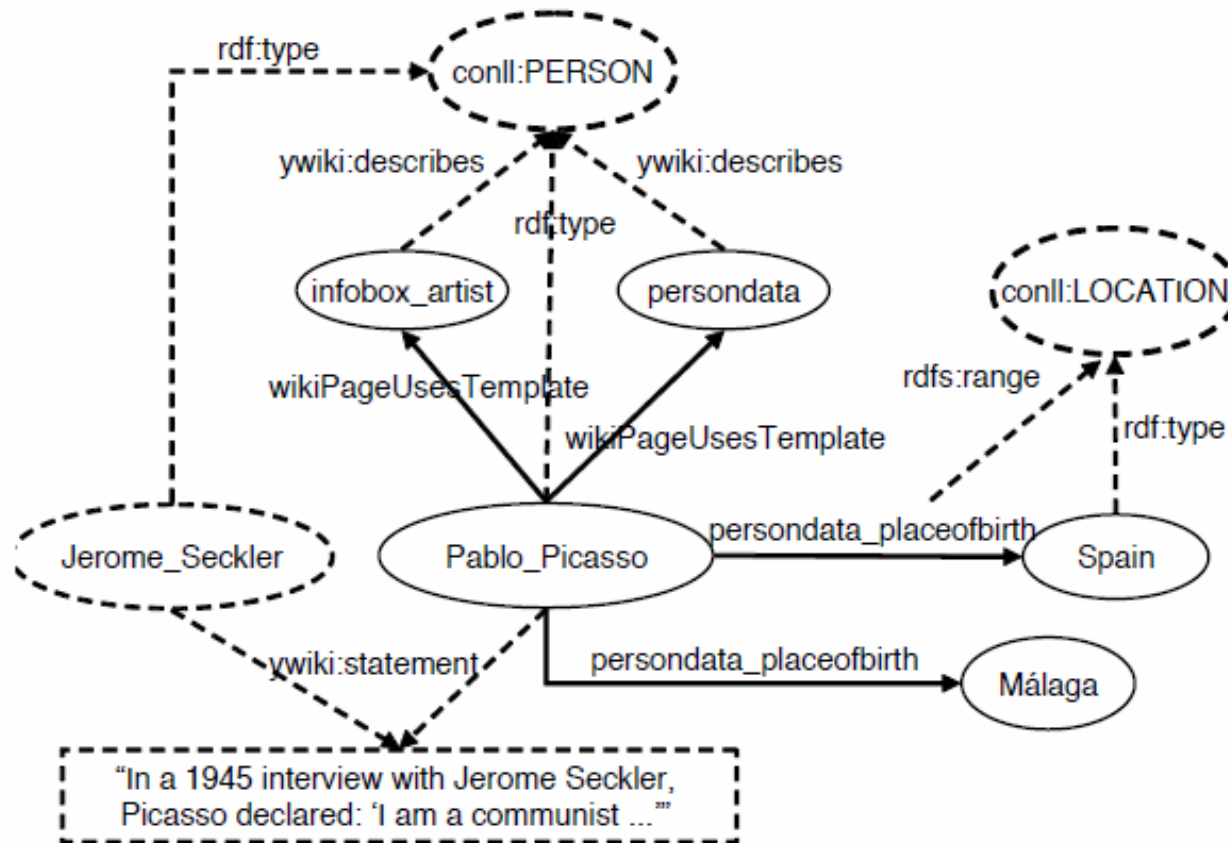


# (property typing examples)

Wikipedia property	CONL L	Examples of property values
infoboxNbaPlayer_name	PER	Alex Groza, Elgin Baylor, Jerry West, David Thompson, Glen Rice, Christian Laettner, Richard Hamilton, Juan Dixon, Sean May, Joakim Noah, ...
infoboxSerialKiller_alias	ORG	Gray Man, the Werewolf of Wysteria, Brooklyn Vampire, Sister, Brian Stewart, Bloody Benders, The Prostishooter, The Rest Stop Killer, The Truck Stop Killer, The Sunset Strip Killer, Cincinnati Strangler, Son of Sam, Plainfield Ghoul, Ed " Psycho" Gein, The Co-ed Killer, ...
highlanderCharacter_born	LOC	Unknown, unknown, 1659, 1945, 802, 1887, 1950, , Glenfinnan, Scotland, (original birth date unknown) ("Highlander II"), 896 BC, Ancient Egypt (original birth date unknown) ("Highlander II"), California, ...
infoboxSuperbowl_stadium	LOC	Sun Devil Stadium, Georgia Dome, Miami Orange Bowl, Hubert H. Humphrey Metrodome, Dolphin Stadium, Raymond James Stadium, Louisiana Superdome, Joe Robbie Stadium, Ford Field, Los Angeles Memorial Coliseum, ...
infoboxWeapon_usedBy	LOC	USA, None, One, none, Italy, United States, Mexico, UK, Russia, Under development, ...
minorLeagueTeam_league	MISC	Eastern League (1923-37, 1940-63, 1967-68, 1992- ), Pacific Coast League, Arizona League, Texas League, South Atlantic League, California League, Midwest League, Northwest League, International League, Carolina League, ...
infoboxTea_teaOrigin	LOC	Nuwara Eliya, Sri Lanka near Adam's Peak between 2200 - 2500 metres, Japan, India, Vietnam, Taiwan, Turkey, China, Anhui, Guangdong, Jiangxi, ...
infoboxPrimeMinister_name	PER	Abdallah El-Yafi, Umar al-Muntasir, Dr. Abdellatif Filali, Abderrahmane Youssoufi, Abdessalam Jalloud, Abdul Ati al-Obeidi, Abdul Hamid al-Bakkoush, Abdul Majid Kubar, Abdul Majid al-Qaud, Abd al-Qadir al-Badri, ...



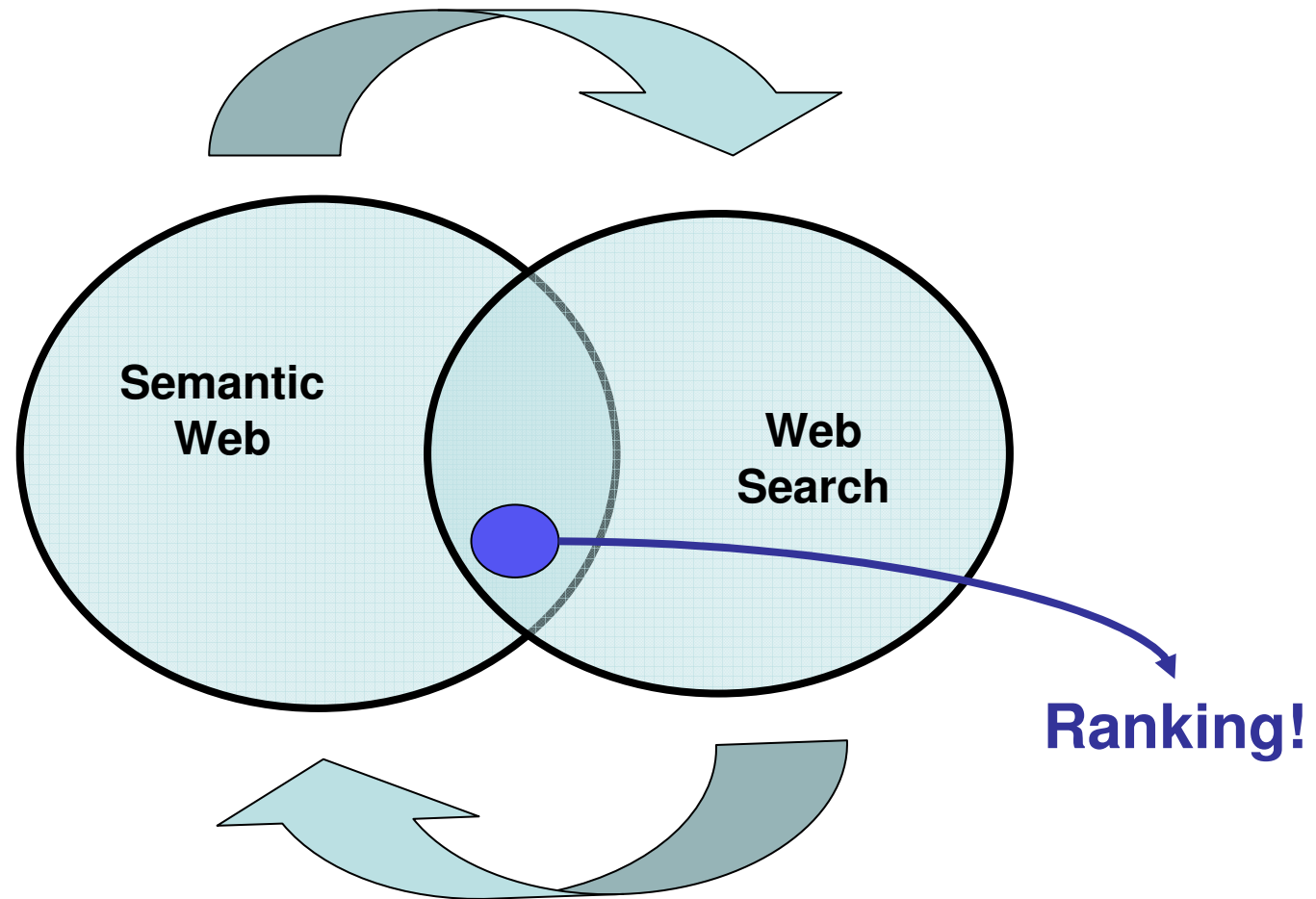
# Beyond annotations, into text



# Plan

- Search Monkey *(today)*
- Web Search: where are we going? *(tomorrow)*
- Current Research Directions *(the future...)*
  - *MicroSearch*
  - *Learning to tag, tagging to learn.*
  - *Ranking Objects*

# Why should you care about search?



# Why is it hard?

- Information needs are intrinsically vague & ambiguous.
- People like to search in natural language.
- Search (Information Retrieval) technology greatly surpassed Boolean querying in the 80s.
- Attempts to improve search technology with semantic knowledge have repeatedly failed since.
  - (except in very narrow domains).
- Effective query expansion is **very difficult**.

# But new directions arise...

- Entity Ranking (sort entities by relevance to a query)
  - Expert Finding
  - Temporal Search
  - Geo-search, ...
- Example:
  - “Effect of sub-prime crisis in Italy”,  
“Pablo Picasso and the Second World War”
    - Find relevant dates
    - Find relevant people & organisations
    - Find relevant geo-locations

# Colored (*Typed*) Indices

Doc #3: The last time **Peter** exercised was in the **XXth** century.

Doc #5: **Hope** claims that in **1994** she run to **Peter Town**.

Peter	→ D3:1, D5:9
Town	→ D5:10
Hope	→ D5:1
1994	→ D5:5
...	
WSJ:PERSON	→ <b>D3:1, D5:1</b>
WSJ:CITY	→ <b>D5:9</b>
WNS:V_DATE	→ <b>D5:5</b>

Possible Queries:

“*Peter* ^ *run*”

“*Peter* ^ **WNS:N\_DATE**”

“(**WSJ:CITY:\***) ^ *run*”

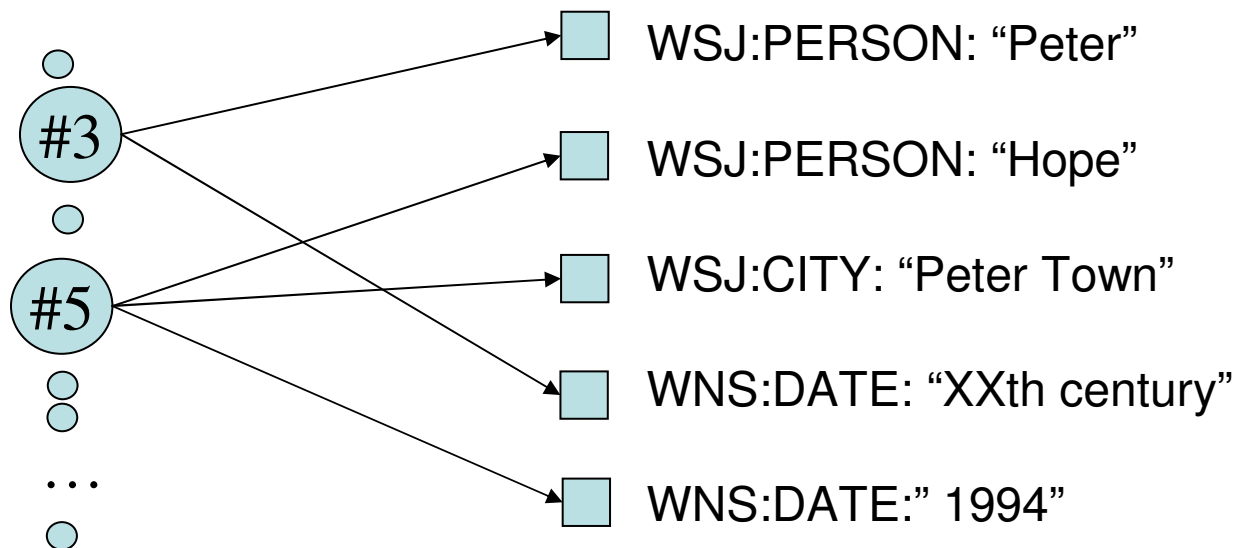
“(**WSJ:PERSON:Hope**) ^ *run*”

*(Bracketing can also be dealt with)*

# Entity containment graphs

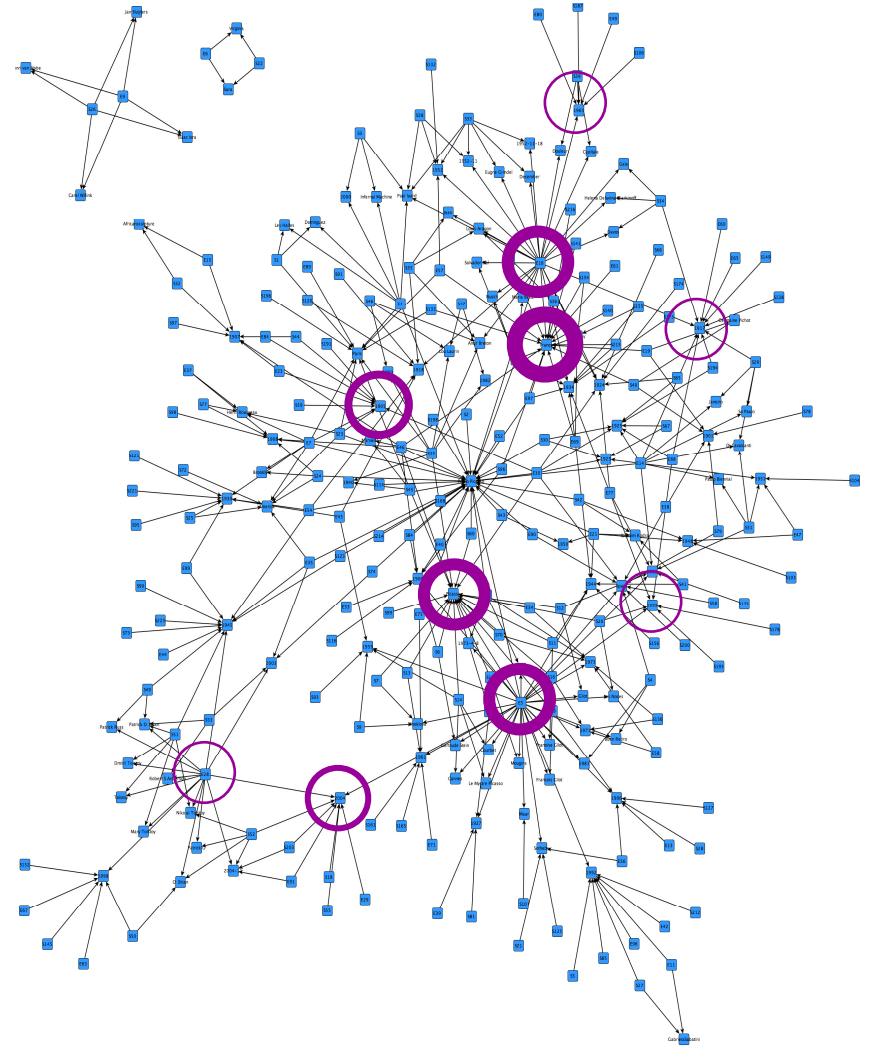
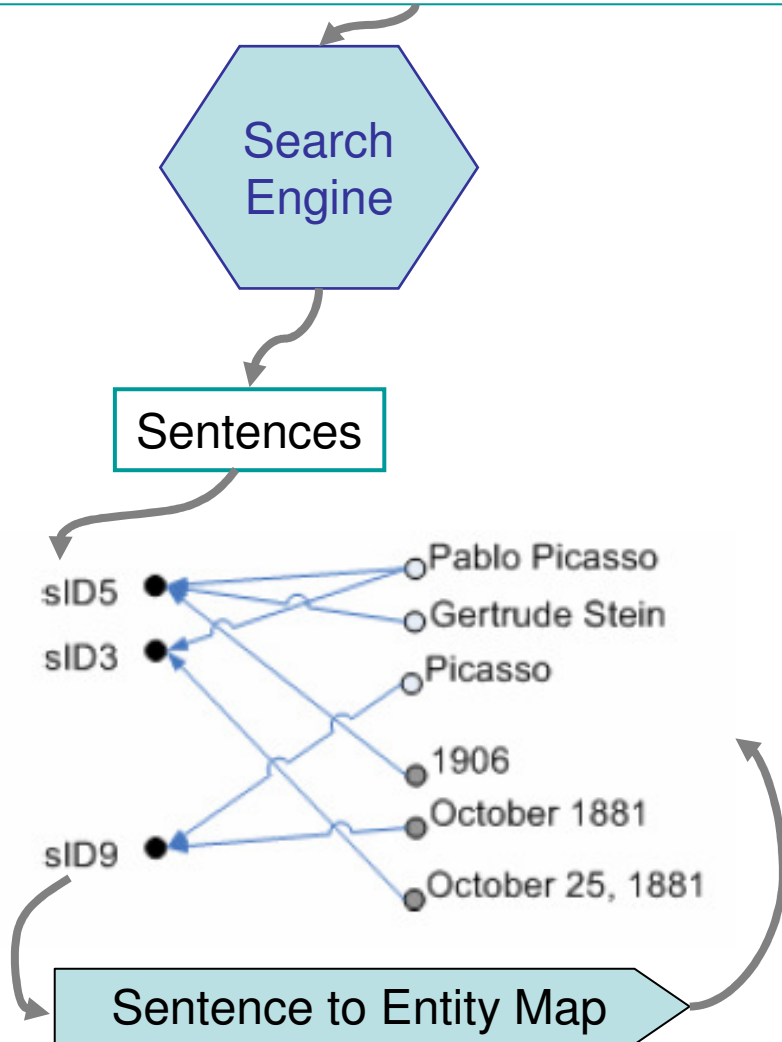
*Doc #3*: The last time **Peter** exercised was in the **XXth century**.

*Doc #5*: **Hope** claims that in **1994** she run to **Peter Town**.



# Entity Conatinment Graph

Pablo Picasso and the Second World War





(Correlator demo)

# Conclusion

- Web Search is moving chaing from Query→Page to:  
Information Need → Information Objects
- A virtuous cycle needs to emerge between Web Search and Semantic Web technologies.
- Difficult scientific problems remain to be solved before we can effectively search Resources:
  - Level of Representation
  - Automatic Extraction and Adaptation
  - Ranking
  - User Interaction