



Searching annotated data.

Hugo Zaragoza

Yahoo! Research (Barcelona)

YAHOO!

- Overall Motivation
- Entity Ranking & Browsing
 - *Yahoo! Reference*
- Exploiting Explicit Annotations
 - *MicroSearch*
 - *InfoBox learning*



Natural Language Retrieval (NLR) Group



Jordi Atserias
- NLP



Mihai Surdeanu
- ML & NLP



Massimiliano Ciaramita
- ML & NLP



Hugo Zaragoza
- ML & IR

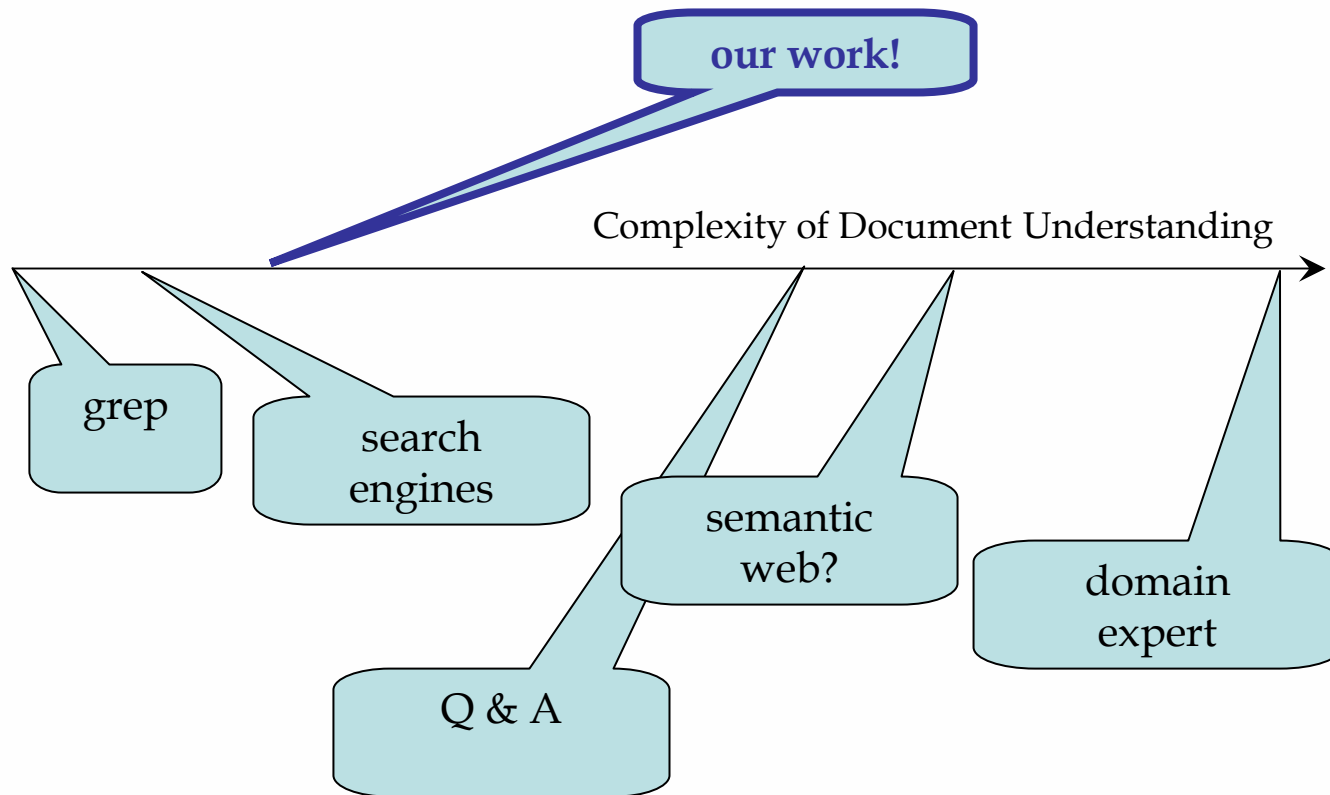


Peter Mika
- Semantic Web

(with help from Giuseppe Attardi, Henning Rode and David Vallet)



Document Understanding Cartoon



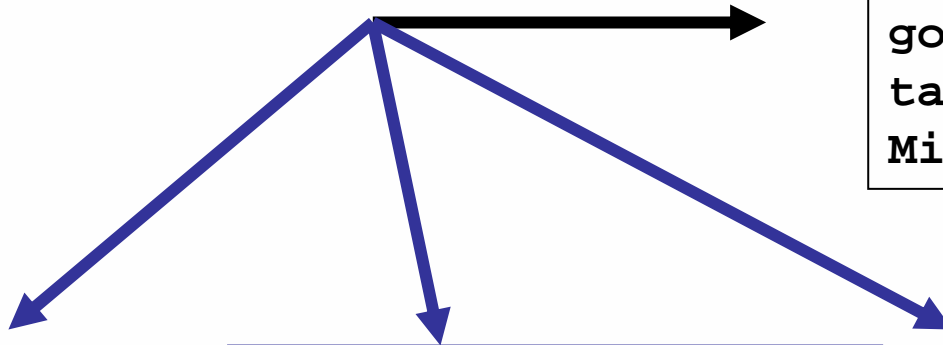


Beyond strings, beyond bag of words

*"In the room the women come and go
Talking of Michelangelo."*



{ room,
women,
come,
go,
talking,
Michelangelo }



{ room ARTIFACT,
women PEOPLE,
come go MOTION,
talking COMMUNICATION,
Michelangelo PERSON }

{ ARTIFACT *room* --MODIFIER,
PEOPLE *women* --SUBJECT ,
MOTION *come and go* --VERB,
}
{
COMMUNICATION *talking* --VERB,
PERSON *Michelangelo* --OBJECT }
}

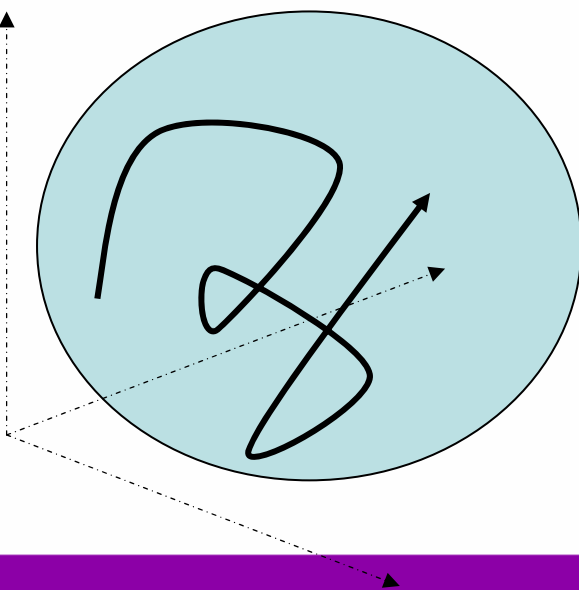
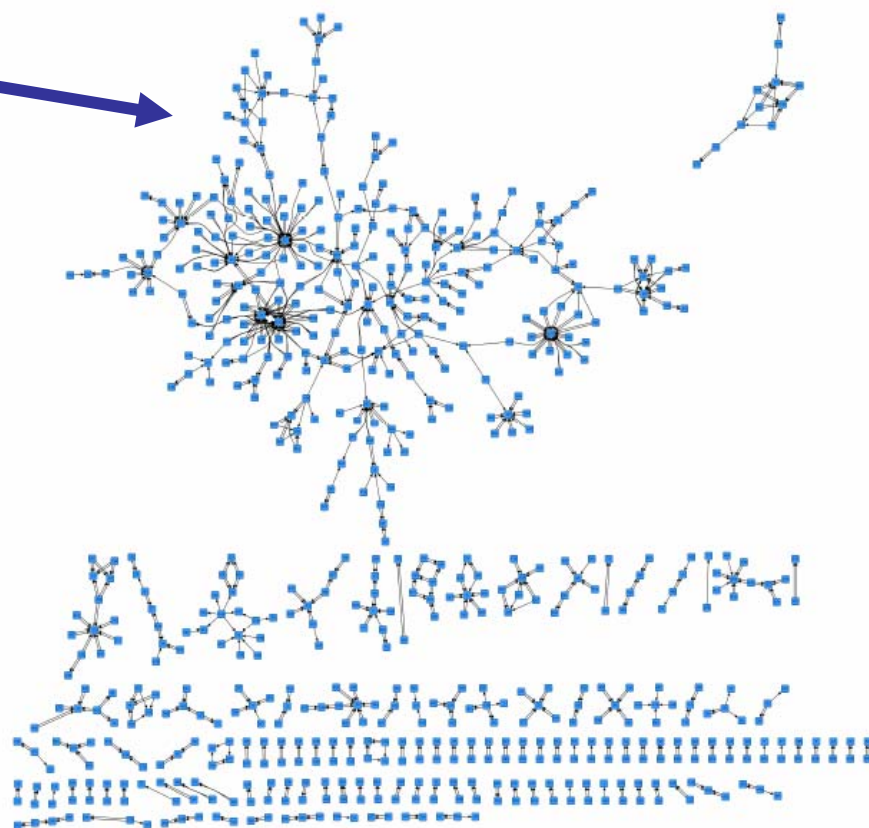
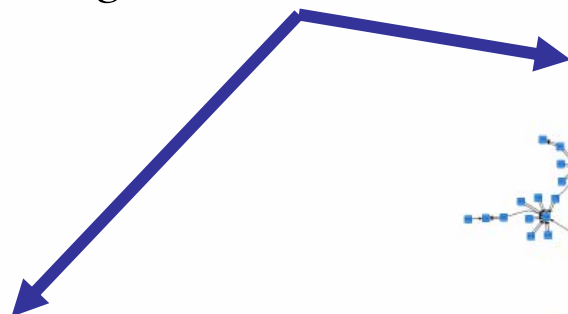
{
PEOPLE-MOTION,
COMMUNICATION-PERSON,
Michelangelo
}

Applications in Search, Advertising, Complex Q&A, ...



Beyond strings, beyond bag of words

*"In the room the women come and go
Talking of Michelangelo."*



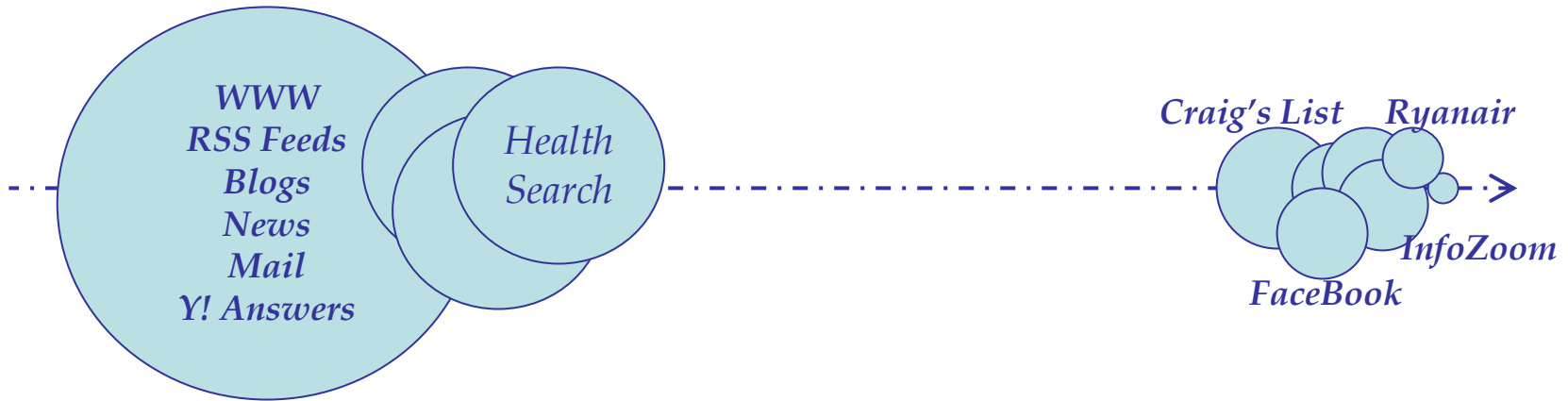


Structure & Domain Knowledge

*Domain
Independent*

*“Verticals”
Specialised Search*

*Domain
Dependent*



Tasks:

Find relevant to string.

Methods:

*String matching (tf-idf)
Hyperlink Popularity*

Tasks:

*AirFlight Booking (find flights)
Housing (find apartment)
Hire People (find CVs)
...*

Methods:

Match + Domain-Based Ranking



Language Representation

Units of representation currently in our system:

- Tokens, Lemmas, POS
- + Semantic Tags
- + Dependency Relations
- + Role labels

Hypothesis:

Linear extraction/parsing time (50K w x m x s)
error-prone output (e.g. 60-90%)
highly redundant information



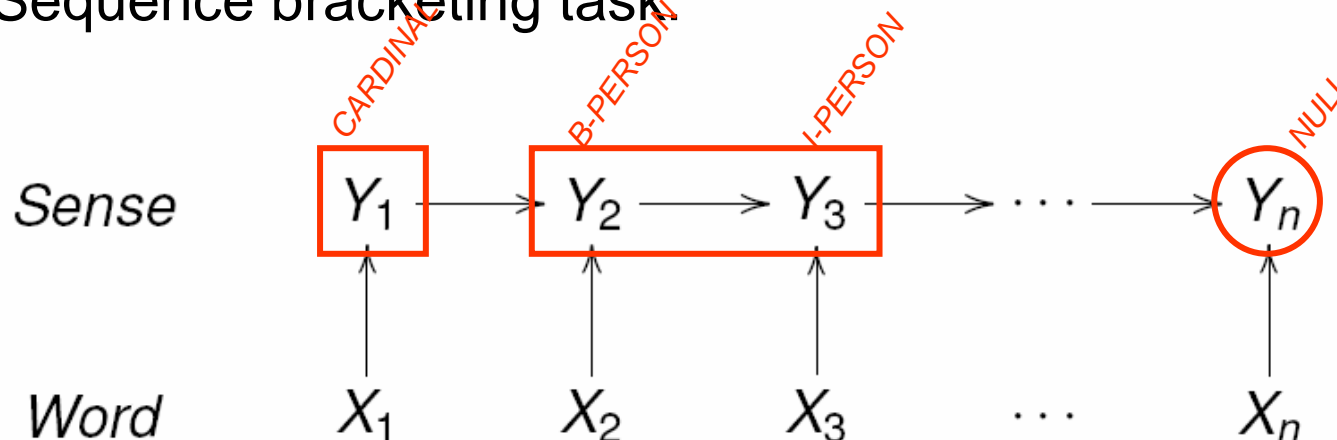
Semantic tagging (some examples)

- Identify instances of semantic **categories** or **relations**:
 - “**Clara Harris**_{person}, one of the guests in the box, stood up and demanded water.” (NER)
 - “**NS-Meg cells**_{cell} expressed **mRNA**_{rna} for the **EPO receptor**_{protein}.” (Bio-NER)
 - “**NS-Meg cells**_{cell} **expressed**_{express_for} **mRNA**_{rna} for the **EPO receptor**_{protein}.” (Relation extraction)
- Broad-coverage/domain-independent
 - “**Clara Harris**_{person}, one of the **guests**_{person} in the **box**_{artifact}, **stood up**_{motion} and **demanded**_{communication} **water**_{substance}.”



Statistical Methods for Semantic Tagging

- Sequence bracketing task:



- Model: Collins Parser (Avg.Perceptron-HMM tagger)
- Features: tokens, POS, word shape, most frequent, previous label, combinations.

(Massimiliano Ciaramita and Yasemin Altun, EMNLP 2006)



(*Publicity Break!*)

<http://www.yrbcn.es/semanticWikipedia>



Applications III: Better Interfaces

Query: Dinosaurs in Argentina

Dinosaur

Dinosaurs were vertebrate animals that dominated terrestrial ecosystems for over 160 million years, first appearing approximately 230 million years ago.

At the end of the Cretaceous Period, approximately 65 million years ago, a catastrophic extinction event ended dinosaurs' dominance on land. [...]



Current dinosaur "hot spots" include southern South America (especially **Argentina**) and China.

Herrerasaurus

Herrerasaurus (meaning "Herrera's lizard," after the name of the rancher who discovered the first fossil of the animal) was one of the earliest carnivorous **dinosaurs**. All known specimens have been discovered in northwest Patagonia, **Argentina**, in late Triassic Period rocks (Early Carnian stage, around 228 million years ago).



Diplodocus

Diplodocus is a genus of diplodocid sauropod dinosaur whose fossilised skeleton was first discovered in 1877 by S. W. Williston. The generic name, coined by Othniel Charles Marsh in 1878, is a Neo-Latin term derived from Greek διπλῶος (diploos) [...]



People



[Philip Currie](#)
[Carlos Munoz](#)
[Charles Darwin](#)
[more \(sociogram view\)...](#)

Other Queries



[Sculpture](#)
[Platonic cylinders](#)
[more \(graph view\)...](#)

Other Sites



[Dinosaurs in Argentina](#)
[\(school.discovery.com...\)](#)
[Huge Dinosaurs in Argentina...](#)
[\(www.msnbc.msn.com...\)](#)
[more](#)

Yahoo! Answers ?!

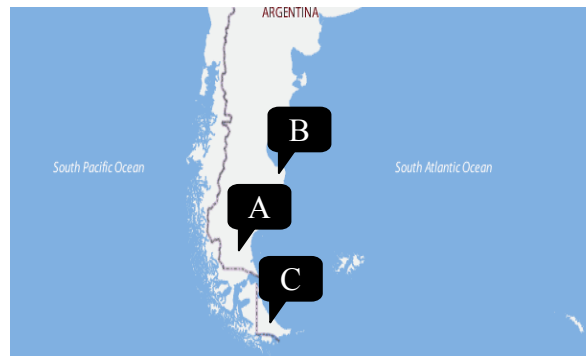
Yahoo! News

Yahoo! Flickr



Places

- A. Neuquen, Argentina
- B. Bariloche Argentina
- C. Ushuaia, Argentina





Applications II: Better Operators

pipes NYTimes through YouTube*

Layout Expand All Collapse All Back to My Pipes New

Sources

User inputs

Operators

- Content Analysis+
- Count+
- Filter+
- Location Extractc+
- Loop+
- Regex+
- Rename+
- Reverse+
- Sort+
- Split+
- Sub-element+
- Tail+
- Truncate+
- Union+
- Unique+

Url

String

Date

Location

Number

Favorites

My pipes

Location Extractor

This module analyzes text within each feed item and attempts to identify addresses, location names or popular map service URLs. If the extractor finds geodata in

Fetch Feed

URL
feed://www.nytimes.com/services/

Content Analysis

Truncate

Truncate feed after 21

Location Extractor

For Each: Replace

Replace each item in input feed with first item output from

[open] YouTube Fetch Data Example

What tag are you searching for at YouTube? item.y:content_an

Pipe Output

- Sources
- User inputs
- Operators
 - Content Analysis+
 - Count+
 - Filter+
 - Location Extractc+
 - Loop+
 - Regex+
 - Rename+
 - Reverse+
 - Sort+
 - Split+
 - Sub-element+
 - Tail+
 - Truncate+
 - Union+
 - Unique+
- Url
- String
- Date
- Location
- Number
- Favorites
- My pipes

Fetch Feed

URL

feed://www.nytimes.com/services/

Entity Extractor +

Type:

- LOCATION
- PEOPLE
- ORGANIZATION
- DATE
- ...

Role:

- SUBJECT
- OBJECT
- MODIFIER
- ...

Dependency Extractor

Anchor:

Target:

- VERB
- SUBJECT
- OBJECT
- MODIFIER
- ...

Sentiment Analysis

- Informational
- Generic/Specific
- Objective/Subjective
- Positive/Negative
- ...

Pipe Output

Location Extractor

This module analyzes text within each feed item and attempts to identify addresses, location names or popular map service URLs. If the extractor finds geodata in the feed, it will annotate each item with a y:location sub-element containing that item's latitude and longitude. Example: [Using the Location Extractor Module](#)

Time taken: 0.16001s [Refresh](#)

- 6 Miners Trapped After Utah Earthquake
- Bomber Kills at Least 28 in Iraq
- Abbas Is Host in Meeting With Olmert
- Duck Signs Up to Hide Duck for Hitting

Debugger: Content Analysis (8



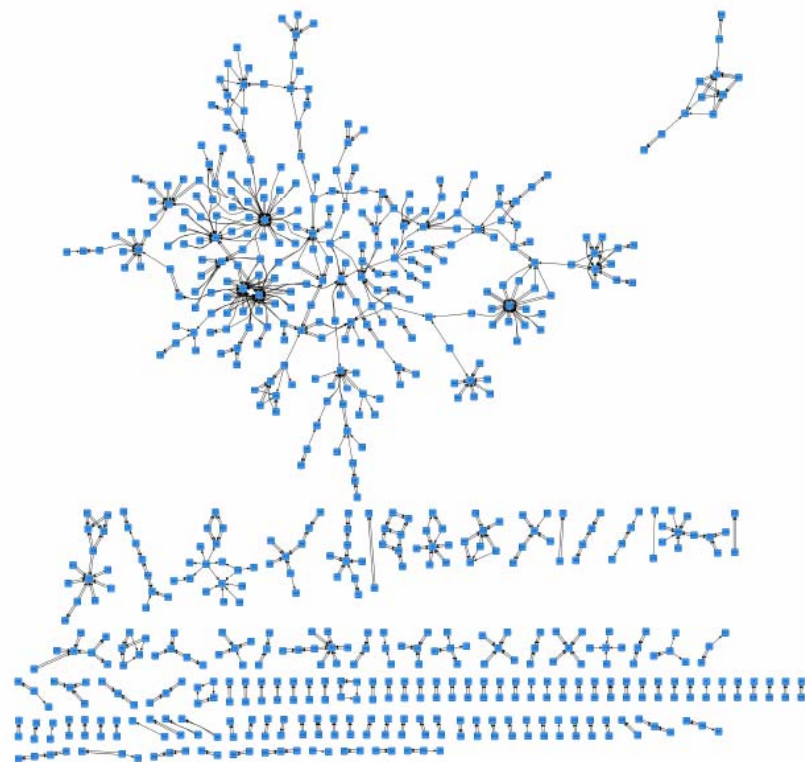
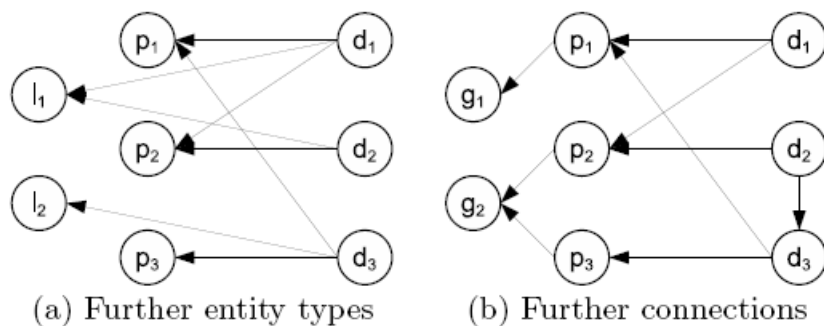
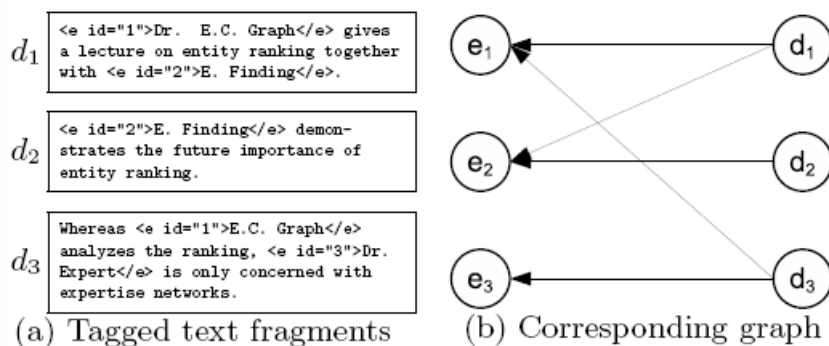
What to do with this data?

- Goal1: better text-features for ranking
 - $d(\text{sentence}, \text{sentence})$
 - $d(\text{question}, \text{answer})$
- Goal2: new units of retrieval
- Goal3: modeling context / intent

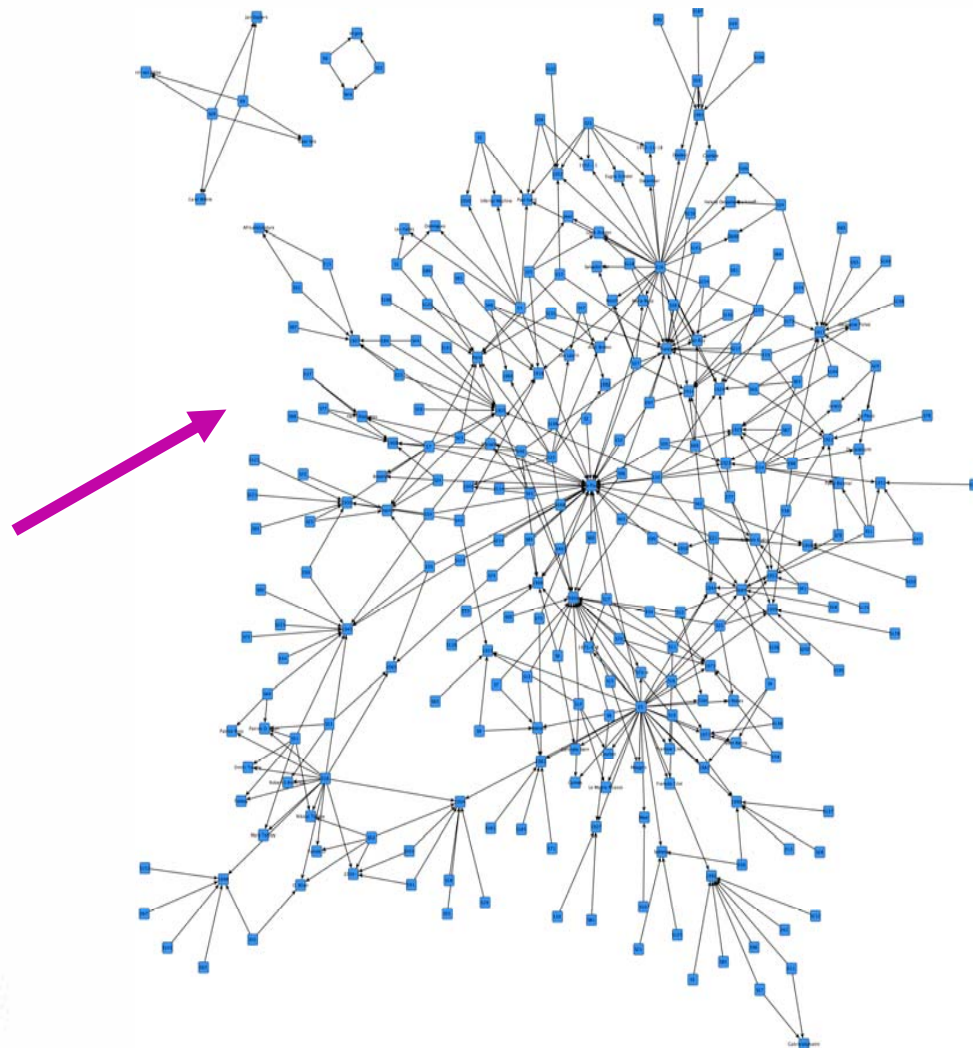
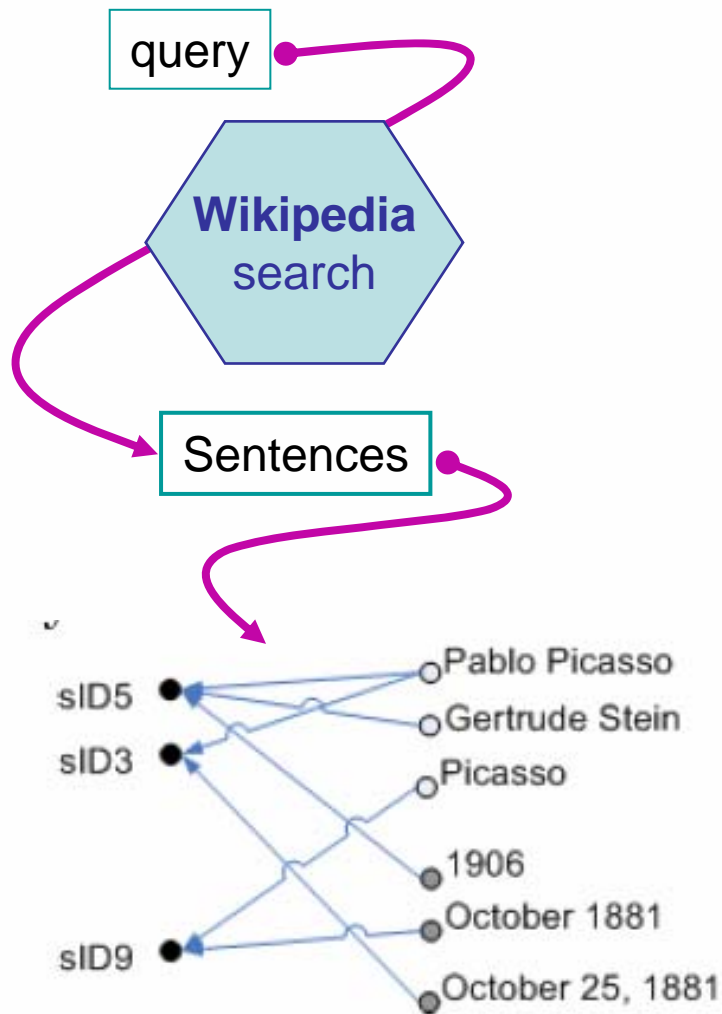


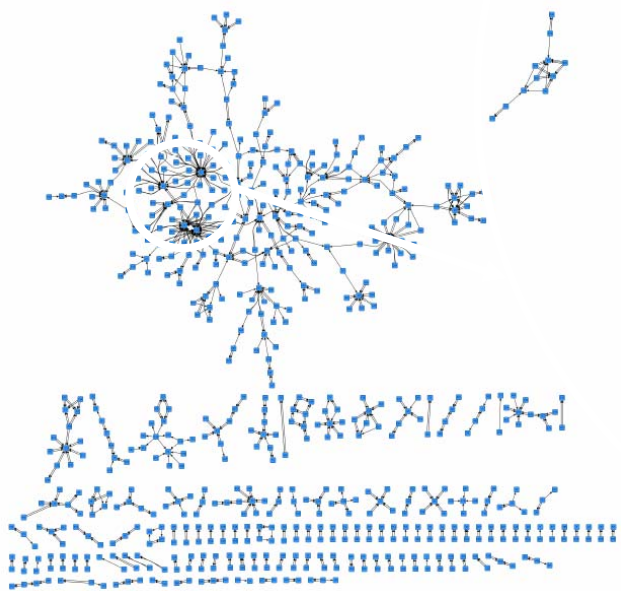
Entity Graphs

- Every Tag (Type) defines an implicit containment graph



[Zaragoza et.al. CIKM07]





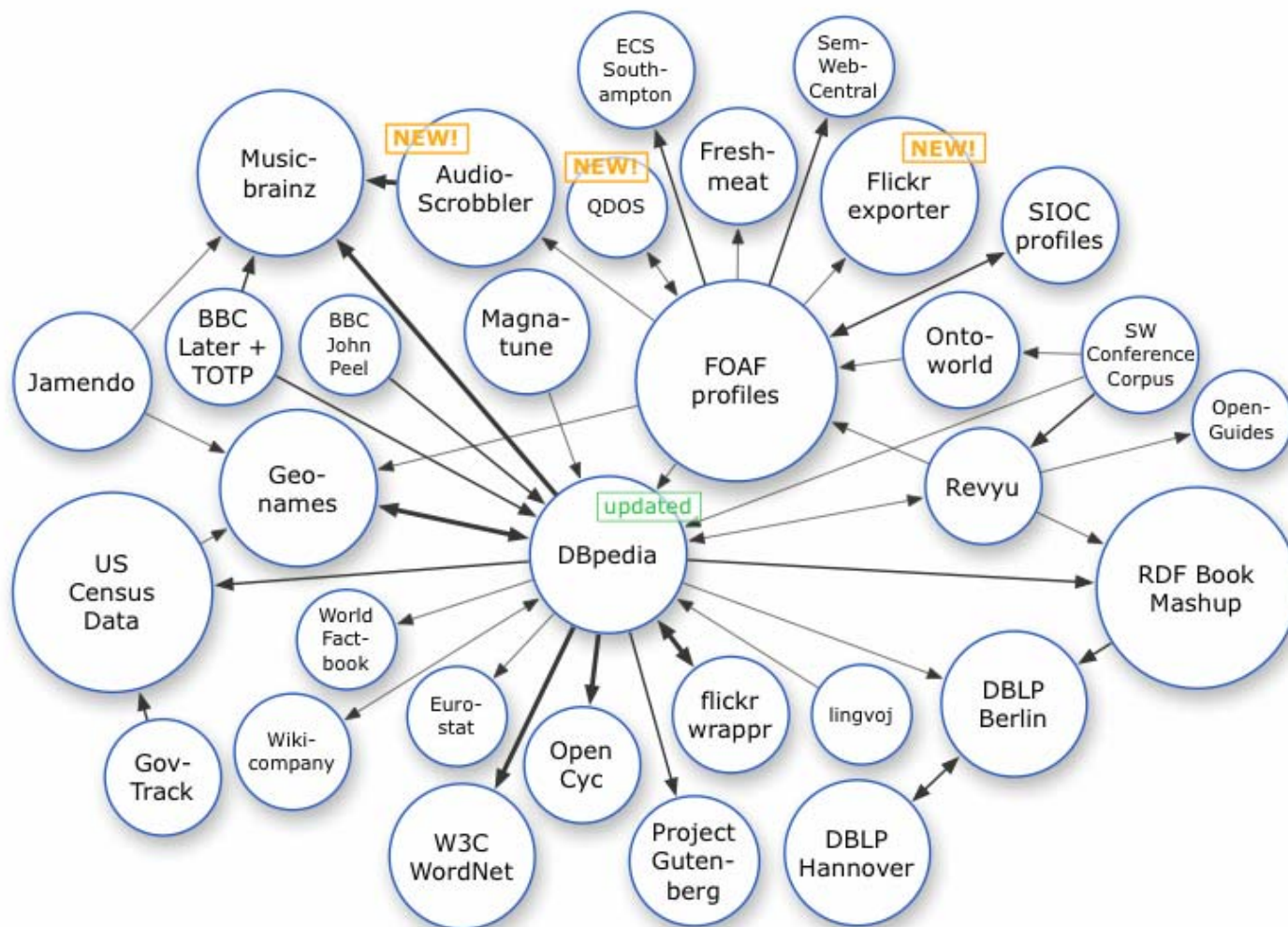


An exciting year for metadata

- Growth of both the Annotated Web and the Data Web
 - Microformats, RDFa
 - Linking Open Data movement
- Opening up of Social Network data and applications
 - Dataportability.org
 - OpenSocial
 - Windows Live Contacts API
- Yahoo!'s Open Search Platform
 - First major search engine to natively support Semantic Web standards
 - Indexing microformat data
 - Part of the Open Search Platform (first out: SearchMonkey)
- Google's Social Graph API
 - Limited to social network data (only XFN and FOAF)



Linking Open Data cloud





Microsearch



ivan herman - Yahoo! Search Results - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://ybcn-svr6.barcelona.corp.yahoo.com:8080/mf/search.do?p=ivan+herman&y=Search&fr=lo

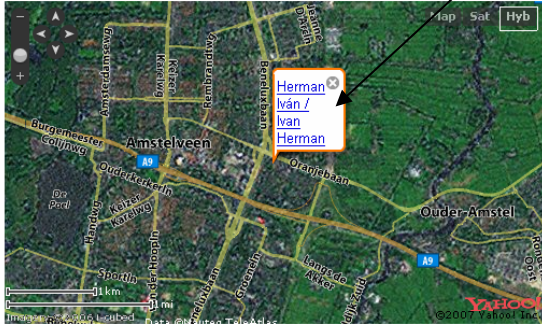
YAHOO! RESEARCH micro SEARCH ivan herman Search

Search Results Persons: 1020 Events: 63 Unfinished: 1 | 1 - 10 of about 289246 for ivan herman - 12 sec. (About this page)

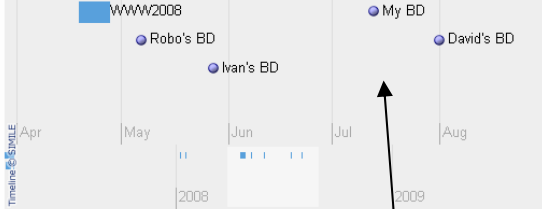
- Ivan's blog**
<http://www.ivan-herman.net/> - 54k [Update metadata](#)
- Ivan Herman**
Ivan Herman gives a talk on behalf of the China Office entitled "What is the ... Ben Adida, Elias Torres, and Ivan Herman give a tutorial entitled "RDFa: ...
<http://www.w3.org/People/ivan/> - 17k [Update metadata](#)
- Ivan Herman's Photos**

name: Herman Iván, Ivan Herman
homepage: <http://www.ivan-herman.net>
My Photos. I like traveling and making photos. ... **Ivãñ Herman**, 2007-12-16,
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/> - 54k [Update metadata](#)
- Ivan Herman**
Ivan Herman. Centrum voor Wiskunde en Informatica (CWI) Kruislaan 413, P.O. Box 94079, ...
Email: Ivan.Herman@cwi.nl or ivan@w3.org. GnuPG: Public GnuPG key ...
<http://homepages.cwi.nl/~ivan/> - 4k [Update metadata](#)
- Ivan Herman - Technical Presentations**
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 - 13k [Update metadata](#)
- ivan herman - ResearchIndex document query**
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**

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/ - 20k [Update metadata](#)
url
[Ivan's blog](#)

Find: yahoo Next Previous Highlight all Match case Done

Geolocation



Related pages based on metadata



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



Example: peter site:flickr.com

Peter site:flickr.com - Yahoo! Search Results - Mozilla Firefox

File Edit View History Bookmarks Yahoo! Tools Help

http://ybcn-svr6.barcelona.corp.yahoo.com:8080/mf/search.do?p=Peter++site%3Aflickr.com&ty=Search&fr=lo

conference san fr... /manager Google Docs & Spr... Elsevier Editorial S... JWS Special Issue ... http://d...38%2F94 Tomcat How to publish Lin... Peter site:flic...

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

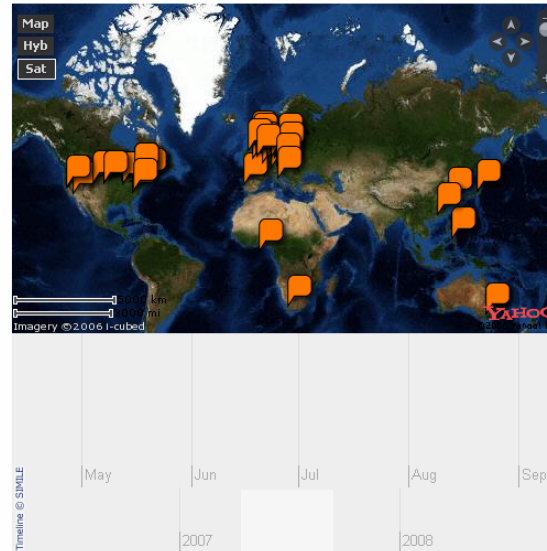
YAHOO! SEARCH Peter site:flickr.com Search Advanced Search

Search Results Persons: 74 Events: 0 Unfinished: 909 | 1 - 10 of about 49209 for Peter site:flickr.com - 10 sec. (About this page)

- Flickr: peter-noster**
Name: Peter Schneider
Address: , MÄ/nster, Germany
Url: <http://blog.peter-noster.de>
Note: I love the movies - I mean I really love them. Film collection: dvd.peter-noster.de/
[Add to your AddressBook](#)
Flickr is almost certainly the best online photo management and sharing ... Testimonials.
peter-noster doesn't have any testimonials yet. You ...
<http://www.flickr.com/people/peter-noster/> - 24k - [Cached](#)
- Flickr: peter bowers**
Name: Peter Bowers
Address: , Toronto, Canada
Note: you can see a picture of me here:
www.flickr.com/photos/pmorgan/5943846/m/photosstream/ and here:
www.flickr.com/photos/pmorgan/44711679/ Nikon D200 Sigma 10-20 mm (this is the wide-angle lens that I take most of my landscape shots with) Nikon Lenses: 20mm f2.8 50mm f1.8 85mm f1.8 35-70mm f2.8 60-200 f2.8 Tamron 90mm f2.8 Macro
[Add to your AddressBook](#)
Flickr is almost certainly the best online photo management and sharing ...
www.flickr.com/photos/pmorgan/44711679/ Nikon D200 ...
http://www.flickr.com/people/mr_fabulous/ - 33k - [Cached](#)
- Flickr: Photos from Peter Ellis**
Name: Peter Ellis <>
Flickr is almost certainly the best online photo management and sharing ... Explore Page Last 7 Days Interesting Calendar A Year Ago Today World Map Camera ...
<http://www.flickr.com/photos/pcellis/> - 25k - [Cached](#)
- Flickr: Peter Kaminski**
Name: Peter Kaminski
Address: , San Francisco, California, USA
Url: <http://peterkaminski.com/>
Note: I'm a friendly and helpful entrepreneur and technologist who lives between San Francisco and Silicon Valley. You may use my photos and images for just about anything, as long as you credit me for the original. The fancy way to say that is that it's licensed under the Creative Commons Attribution license. If you can give me credit for the photo somehow, you don't even have to ask permission, just go ahead and use it. I do love to hear where my photos get to, though; if you can, drop me a line here, or at kaminski@istori.com, or to: Peter Kaminski 655 High Street Palo Alto, California 94301 USA
[Add to your AddressBook](#)
Flickr is almost certainly the best online photo management and sharing ... **Peter Kaminski.** 655

Find: native Next Previous Highlight all Match case

Done Proxy: None



Flickr users named
“Peter” by
geography



Example: san francisco conference

san francisco conference - Yahoo! Search Results - Mozilla Firefox

File Edit View History Bookmarks Yahoo! Tools Help

http://ybcn-svr6.barcelona.corp.yahoo.com:8080/mf/search.do?p=san+francisco+conference&y=Search&fr=lo

virtuosos transaction deadlocked

conference san... /manager Google Docs & ... Elsevier Editori... JWS Special Iss... Problem loading... Tomcat How to publish ... san francis... Recognize Micro...

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

YAHOO! SEARCH Web Images Video Local Shopping more »

san francisco conference Search Advanced Search

Search Results Persons: 20 Events: 135 Unfinished: 1031 | 1 - 10 of about 2725643 for san francisco conference - 10 sec. (About this page)

- San Francisco Conference -- Encyclopaedia Britannica**
Name: Robert McHenry
San Francisco Conference: (April 25-June 26, 1945), international meeting that ... Stalin selected **San Francisco** as the site of a **conference** to take place ...
<http://www.britannica.com/eb/article-9065303/San-Francisco-Conference> - 57k - [Cached](#)
- Oracle OpenWorld - San Francisco**
OpenWorld is the premier Oracle event for business decision-makers, IT ... | Worldwide PRODUCTS AND SERVICES INDUSTRIES SUPPORT PARTNERS COMMUNITIES ABOUT ...
<http://www.oracle.com/openworld/sanfrancisco/conference/index.html> - 36k - [Cached](#)
- ad:tech Interactive Media Conference - San Francisco Events**
ad:tech Interactive Media **Conference** hosted in **San Francisco**. ... **Conference:** Expo Hall: Tuesday, April 15. Wednesday, April 16. Thursday, April 17 ...
<http://www.ad-tech.com/sl/> - 14k - [Cached](#)
- Only In San Francisco - The Official Visitors Site for San Francisco**
Only in **San Francisco** is The Official Visitors Site for **San Francisco**. ... through the square, and you'll see **San Francisco** in all its melting-pot glory. ...
<http://onlyin.sfvistor.org/> - 53k - [Cached](#)
- Gilbane San Francisco 2008**
Gilbane Content Management Technologies **conference** at the Palace Hotel **San Francisco**, June 18 - 20, 2008 information. ... **Conference** Events. Gilbane **San** ...
<http://gilbanesf.com/> - 12k - [Cached](#)
- South San Francisco Conference Center**
Choice for your next product launch, **conference**, convention, trade show, seminar, reception or banquet.
<http://www.ssfconf.com/> - 14k - [Cached](#)
- 5th Annual Conference on Arteriosclerosis, Thrombosis and Vascular Biology**
... Hilton **San Francisco** - **San Francisco**, CA. **Conference** Dates: May ... rooms at the Hilton **San Francisco** for **conference** participants and their guests. ...
<http://www.americanheart.org/presenter.jhtml?identifier=3015499> - 41k - [Cached](#)
- ISPI Annual Conference 2007, San Francisco, California**
Performance Improvement **Conference**. Performance Beyond Borders. **San Francisco** Marriott Hotel ... Come to **San Francisco**, an international city alive with ...
<http://www.ispi.org/ac2007/> - 17k - [Cached](#)
- Treaty of San Francisco - Wikipedia, the free encyclopedia**
were invited to the **San Francisco Peace Conference**, and therefore neither signed this treaty.

Find: native Next Previous Highlight all Match case

Done Proxy: None

Map
Hyb
Sat

0 1000 km
0 1000 mi

Imagery © 2006 empty

2007
@media 2007 America
Workshop: Transcending CSS
Open Source Convention
@media 2007 Europe
Future of Web Design Europe
Bu Pri

Timeline © SHALE

Imagery © 2006 empty

2007 2008

Conferences in San Francisco date



Example: greater st. peter

greater st.peter - Yahoo! Search Results - Mozilla Firefox

File Edit View History Bookmarks Yahoo! Tools Help

http://ybcn-svr6.barcelona.corp.yahoo.com:8080/mf/search.do?p=greater+st.peter&y=Search&fr=lo

conference san fr... /manager Google Docs & Spr... Elsevier Editorial S... JWS Special Issue ... http://d...38%2F94 Tomcat How to publish Lin... greater st.pe...

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

Web Images Video Local Shopping more »

YAHOO! SEARCH greater st.peter Search Advanced Search

Search Results

Persons: 18 Events: 1 Unfinished: 531 | 1 - 10 of about 897807 for greater st.peter - 13 sec. (About this page)

- 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
- 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
- Greater Saint Pete**
Christian congregation
http://www.poweredbyc
- St Peter's RC Schc**
... reach for greater dn
Buffalo. Campus Outfit
http://www.stpeterc.on
- Jobs at The YMCA**
Jobs at The YMCA of C
engine. ... ymca of gre
http://www.careerjet.co
- St. Peter's Episcop**
The Homepage of St. P
community, to work to
http://www.stpeterschic
- St. Louis Honda de**
St. Louis Honda dealer
new Honda & used Hor
http://www.stpetershon
- Greater St Peter E**
Get details on Greater
user reviews & editorial
http://www.citysearch.c

Map Hyb Sat

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

Web Images Video Local Shopping more »

YAHOO! SEARCH greater st.peter Search Advan

Search Results

- 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

Find: native Done

Call on Skype

Add to your addressbook

(other actions)



Bridging implicit and explicit metadata.

- 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**.^[1] 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.^[2] 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,^[3] his first word was "piz," a shortening of *lápiz*, the Spanish word for [pencil](#).^[4] 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](#)



Result #1: extending metadata

Pablo Picasso was born in Málaga, Spain.

PER

LOC

LOC

E:PERSON

GPE:CITY

GPE:COUNTRY

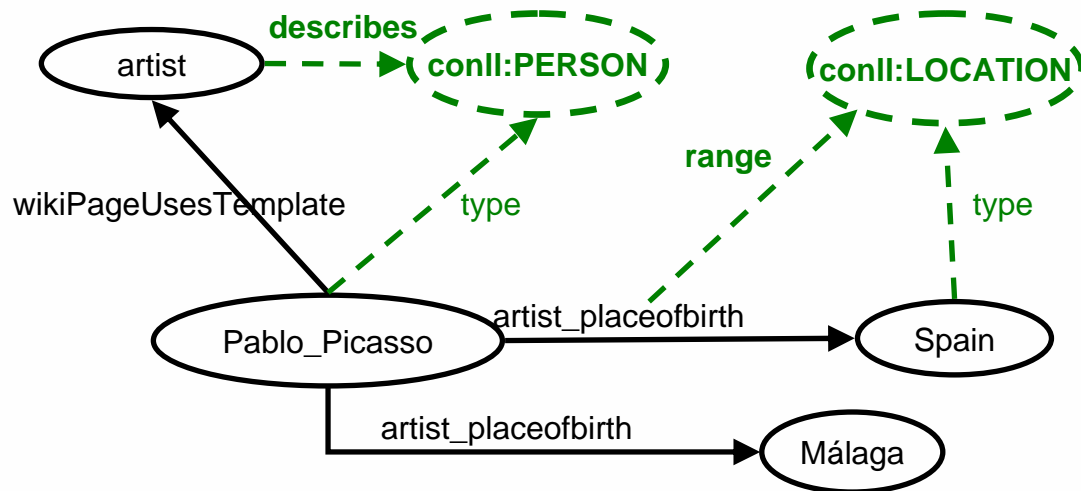
artist:name

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.





Examples

	PER	ORG	LOC	MISC	Σ
rdfs:range	805	515	480	124	1924
ywiki:describes	184	18	13	4	219
rdf:type (from text)	285873	140885	46079	32550	505387
rdf:type (through rdfs:range)	307874	58768	31383	8108	406133
rdf:type (through ywiki:describes)	47702	489	611	74	48876

Wikipedia property	CONLL	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 Yousoufi, Abdessalam Jalloud, Abdul Ati al-Obeidi, Abdul Hamid al-Bakkoush, Abdul Majid Kubar, Abdul Majid al-Qaud, Abd al-Qadir al-Badri, ...



Some results

- Part of what we learn is a mapping between tagsets
- Generate training data based on what is learned
 - e.g. if Paris is also a place of birth, it is also a place
- Dealing with sparse data

Paris is the capital city of France.

- As good as manually created training data!
- Improvement when combined
 - On news: +6.1%
 - On Wikipedia: +5.5%



Trying to find answers to complex questions...

Q: How do you quiet a squeaky door?

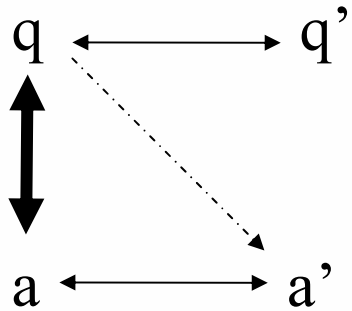
A: Spray WD-40 directly onto the hinges of the door. Open and close the door several times. Remove hinges if the door still squeaks. Remove any rust, dirt or loose paint. Apply WD-40 to removed hinges. Put the hinges back, open and close door several times again.

Q: How do I open up more and become not so shy around woman?

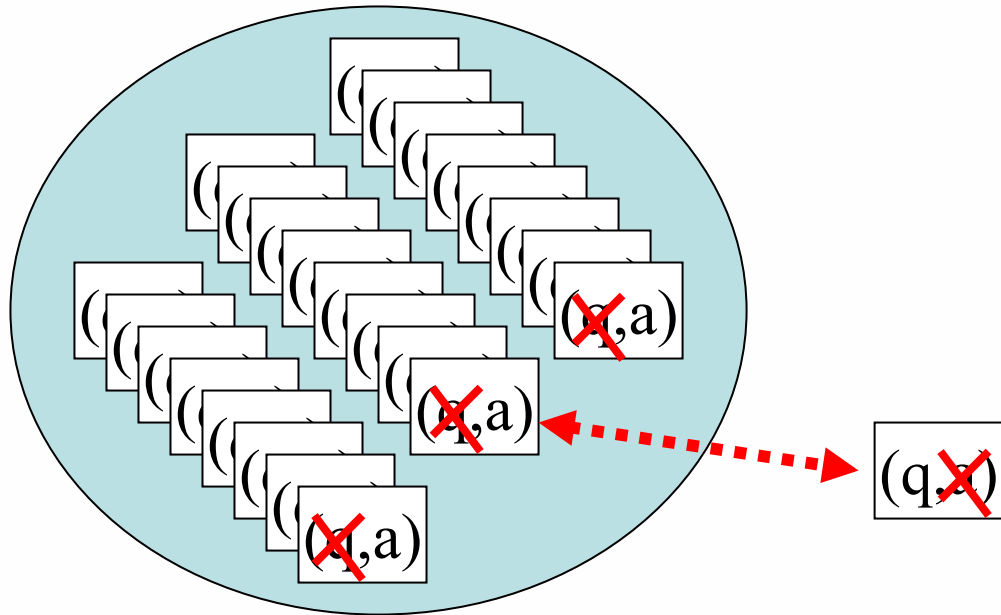
A: Sorry, me too!!



Trying to find answers to complex questions...



$d(q', s)$
 $d(q', q)$
 $d(q', q, a)$
 $d(q', a)$





Thank you!
