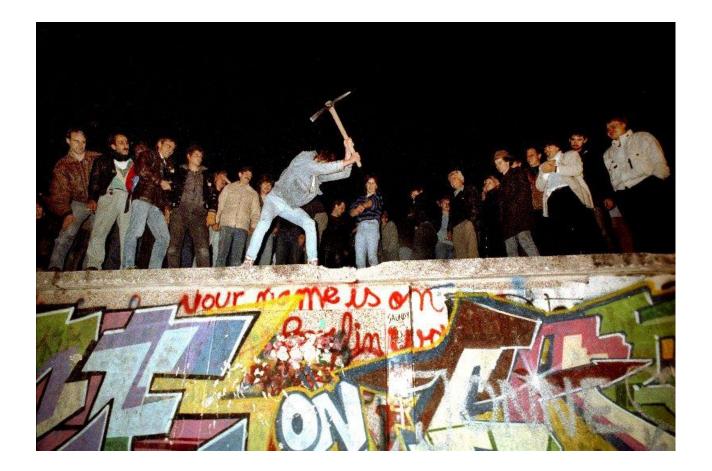
CC7220-1 LA WEB DE DATOS PRIMAVERA 2022

LECTURE 1: INTRODUCTION

Aidan Hogan aidhog@gmail.com



THE WEB IS NOW 3 DECADES OLD



The future of the Web?



What will the Web be like in 3 decades time?







The Semantic Web?

Semantic Web?



semantic web

Google Search

I'm Feeling Lucky

Ŷ

Semantic Web?



"The Semantic Web will bring structure to the meaningful content of Web pages, creating an environment where software agents roaming from page to page can readily carry out sophisticated tasks for users."

– Berners-Lee et al. (2001) "The

Sci. American

Consider answering: "What is the Web?"

What's wrong with the current Web?

THE CURRENT WEB IS GREAT!

← → C 🗋 sometimesedsometimesblue.com	\$ 0	≡

THE CURRENT WEB IS GREAT!



Fifty years ago, if you wanted to see a picture of a raccoon, you either had to already have it or drive to a library. And raccoon in a funny hat? Forget about it.

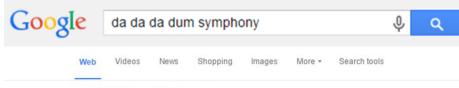
Google is also pretty great

Google

movie what's in the box

Seven (stylized as SE7EN) is a 1995 American crime thriller film directed by David Fincher and written by Andrew Kevin Walker. It stars Brad Pitt, Morgan Freeman, Gwyneth Paltrow, John C. McGinley, R. Lee Ermey, and Kevin Spacey.

Seven (1995 film) - Wikipedia https://en.wikipedia.org/wiki/Seven_(1995_film)



About 107,000 results (0.36 seconds)



Beethoven - Symphony No. 5 in C Minor (1) - YouTube www.youtube.com/watch?v=W2qW6fOtAMY *



sometimes w	hen i'm			
sometimes w	hen i'm alone i use	e comic sans		
sometimes w	hen i'm alone i go	ogle myself		
sometimes w	hen i'm alone i cry	,		
sometimes w	hen i'm all alone			
sometimes w	hen i'm dreaming			
sometimes when i'm sad i like to cut myself				
sometimes w	hen i'm dreaming	lyrics		
sometimes w	hen i'm alone			
sometimes w	hen i'm driving on	the road at night		
sometimes w	hen i'm <mark>alone i wo</mark>	nder		
	Google Search	I'm Feeling Lucky		



J

Q

www.youtube.com

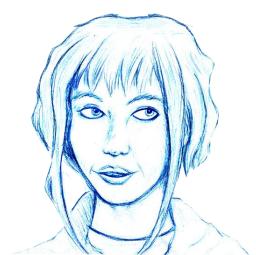
PROBLEM WITH THE CURRENT WEB: LITERATURE VETERANS

DOING A REPORT FOR UNIVERSITY ...



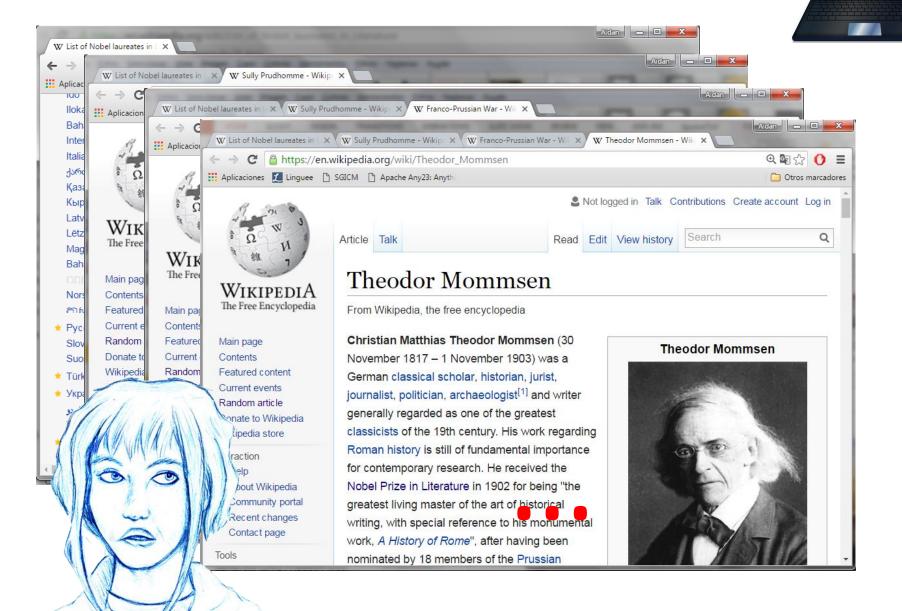
Wants to find all:

- Nobel Prize winners in Literature
- Who fought in at least one war
- The year they won the prize
- And the year the war(s) started



... how would you solve this?

Lots of Wikipedia Tabs ...



The Laureate Algorithm (3 decades on)



results := Ø

for all nobel-lit-winner in wiki-list

```
year-prize := year of nobel-lit-winner
```

```
wars := search "war", "conflict", "battle" in nobel-lit-winner
```

```
for all war in wars
```

```
year-war := starting year of war
```

```
add nobel-lit-winner, year-prize, war, year-war to results
```

end

end

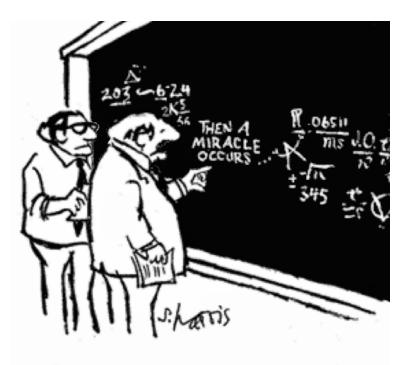
return results



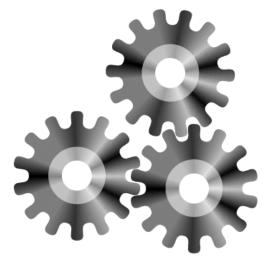




nobel-lit-winner-and-wars := <u>magical-sem-web-results()</u> return nobel-lit-winner-and-wars



"I THINK YOU SHOULD BE MORE EXPLICIT HERE IN STEP TWO,"

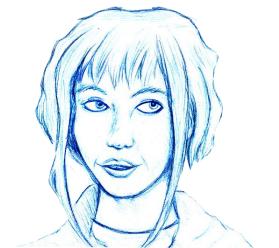


DOING A REPORT FOR UNIVERSITY ...



Wants to find all:

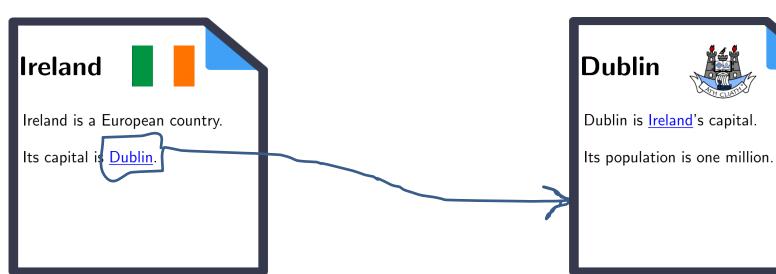
- Nobel Prize winners in Literature
- Who fought in at least one war
- The year they won the prize
- And the year the war(s) started



... why is this query hard on the current Web?

So what's the problem ...

THE CURRENT WEB IS DOCUMENT-CENTRIC



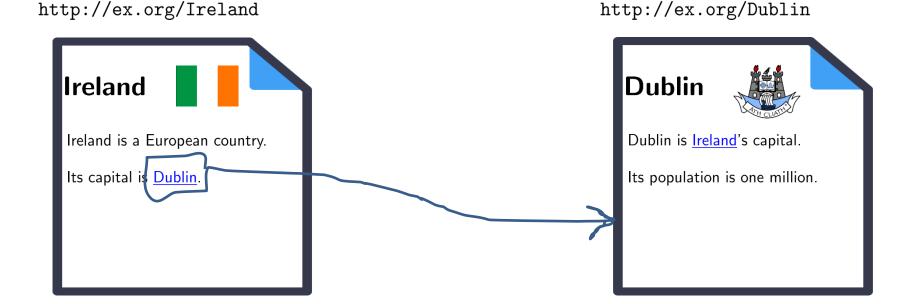
http://ex.org/Dublin

<html>

http://ex.org/Ireland

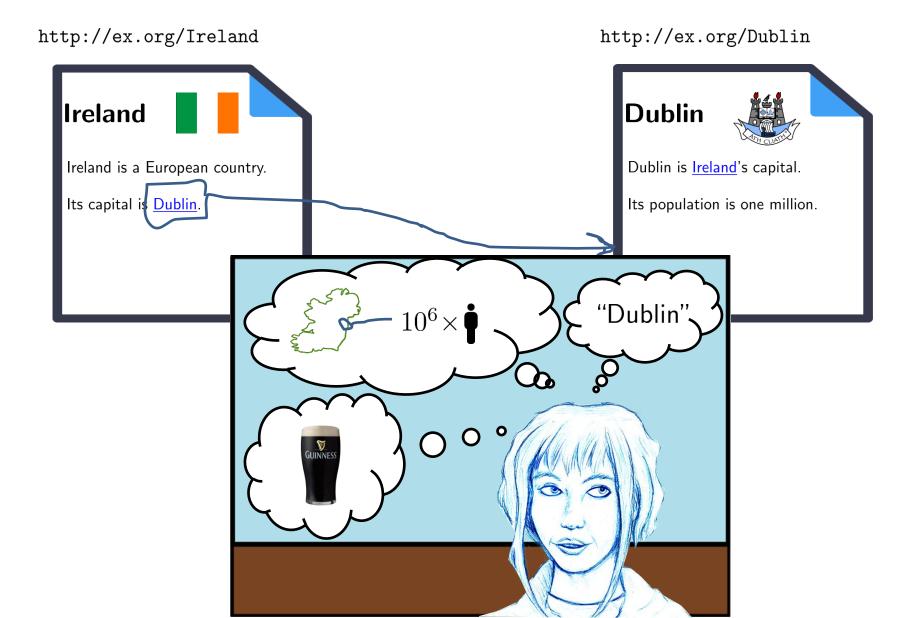
</html>

THE CURRENT WEB IS DOCUMENT-CENTRIC

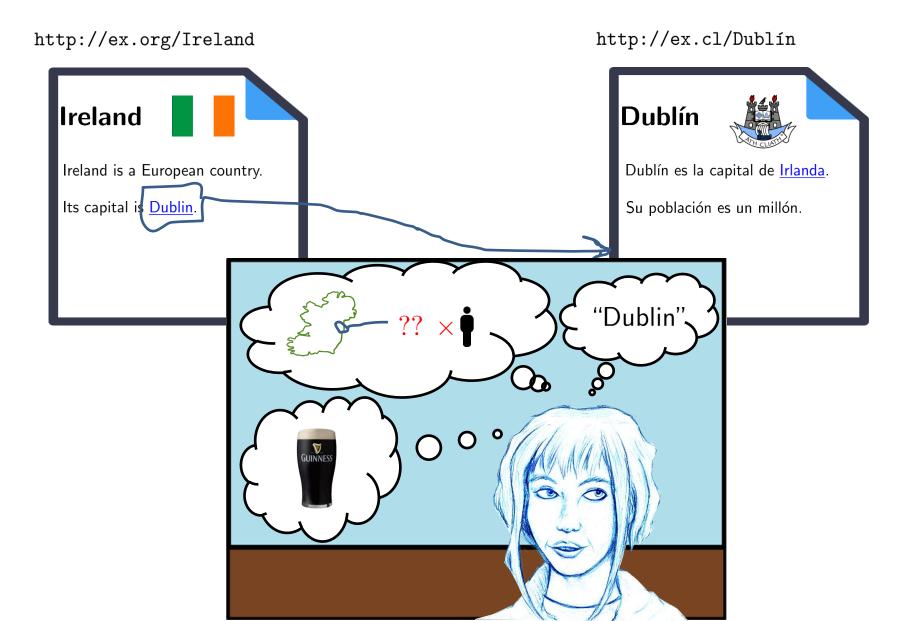


```
<html>
        <body>
        <h1>Ireland</h1>
        <div class="flag"><img src="flag.jpg" /></div>
        Ireland is a European country.
        Its capital is <a href="http://ex.org/Dublin">Dublin</a>.
        ...
        </body>
        </html>
```

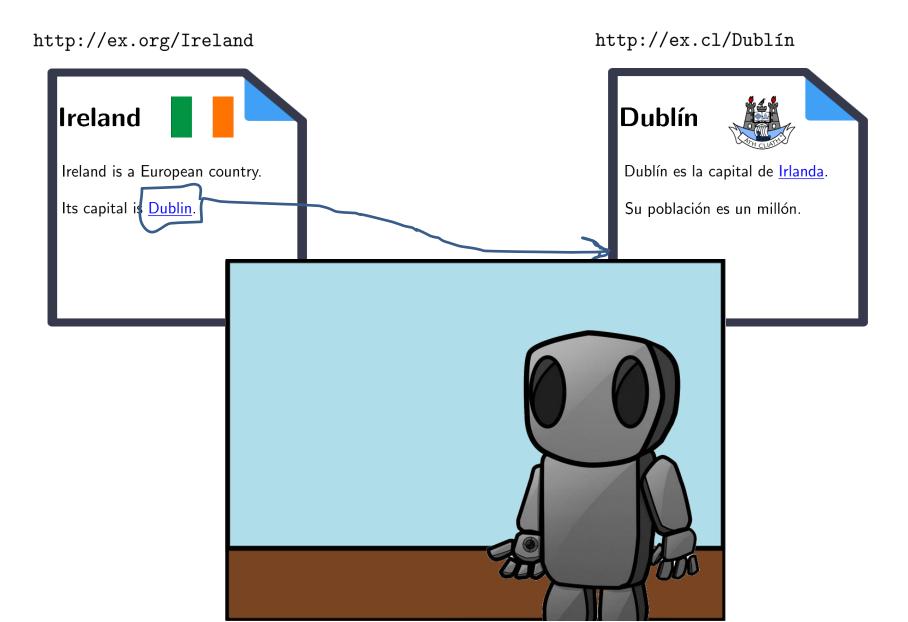
(Most of it) Makes sense to humans



... ASSUMING THEY SPEAK THE LANGUAGE



EVEN WORSE FOR MACHINES



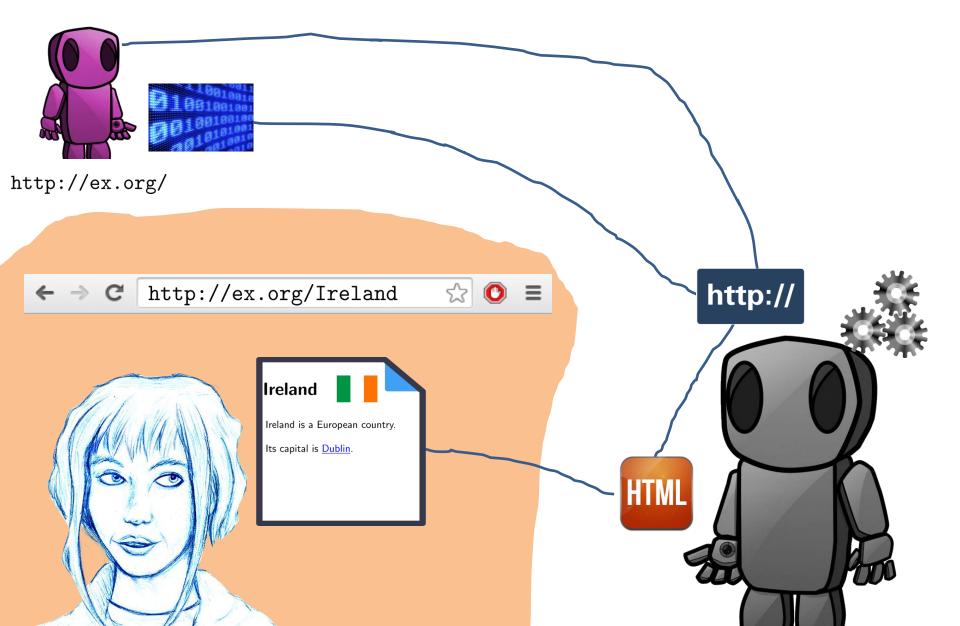
YTYL NUGZY PUG MEBHOLYZ

http://ex.org/Ogyselw Ogyselw Wdcsól Ogyselw oz e Ydguryel budlfgi. Wdcsól yz se berofes wy Ogselwe. Ofz berofes oz Wdcsol Zd rucseboúl yz dl mossúl. 0 0 0

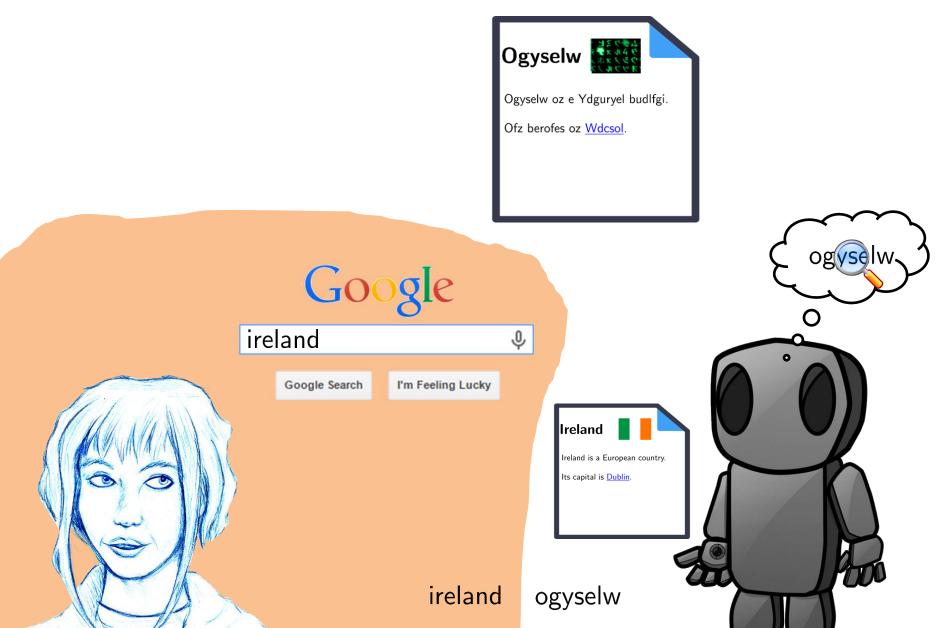
http://ex.cl/Wdcsól

So what's the problem ...

What machines can do: Fetch documents



What machines can do: Find documents



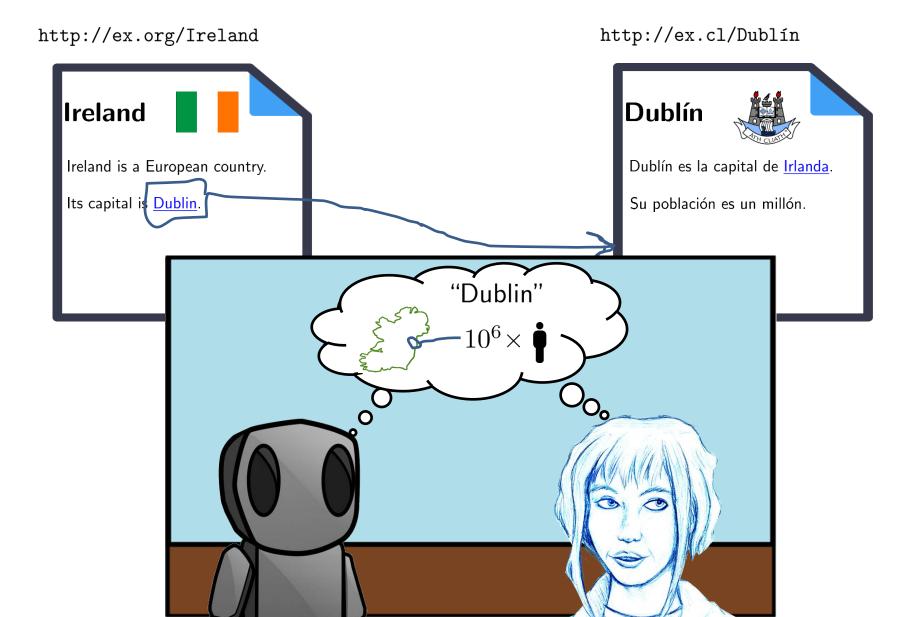
What machines cannot do: Combine sources

http://ex.org/Ogyselw Ogyselw Wdcsól Ogyselw oz e Ydguryel budlfgi. Wdcsól yz se berofes wy Ogselwe. Ofz berofes oz Wdcsol Zd rucseboúl yz dl mossúl. 0 0 0

http://ex.cl/Wdcsól

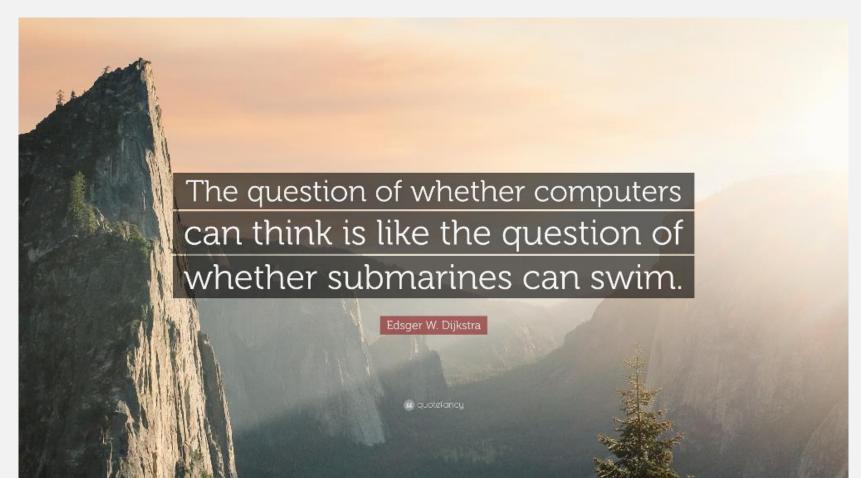
TOWARDS A SEMANTIC WEB

MACHINES THAT "UNDERSTAND" THE WEB?

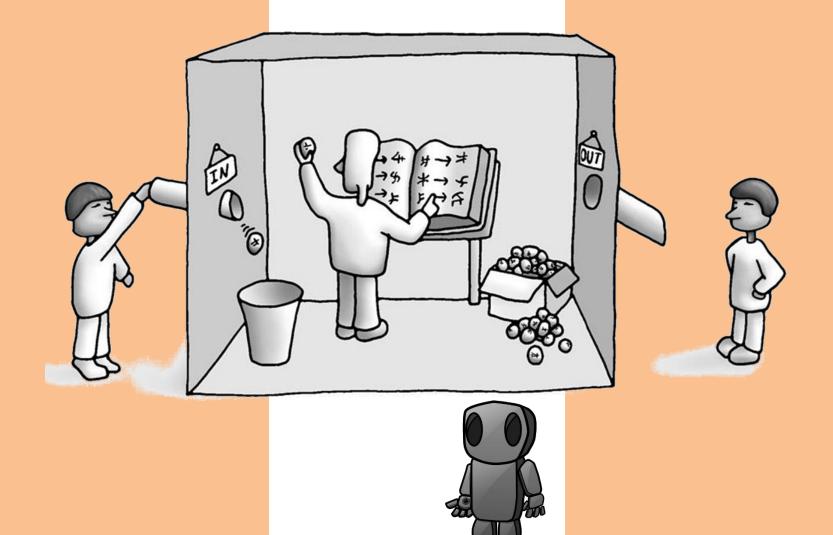


Edsger W. Dijkstra



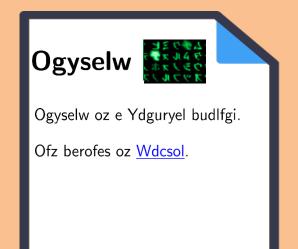


SEARLE'S CHINESE ROOM ...



SEARLE'S CHINESE ROOM (NATURAL LANGUAGE)

http://ex.org/Ogyselw



http://ex.org/Wdcsol



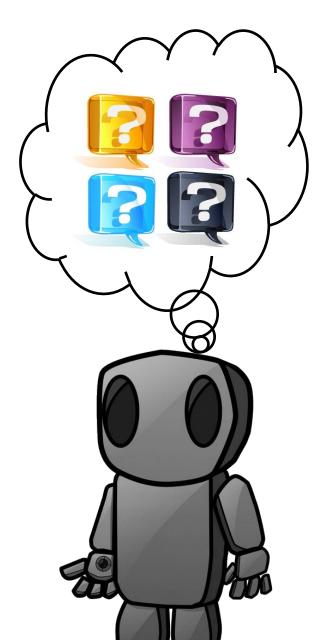
INPUT: "Nhef oz fhy rurdsefoul up fhy berofes up Ogyselw?"

... what should the output be?

 $OUTPUT: \ ``uly \ mossoul''$

MULTIPLE NAMES, ONE THING ...





ONE NAME, MULTIPLE THINGS ...





MULTIPLE WAYS TO SAY THE SAME THING ...

Dublin's population is one million. Dublin has a population of one million. Dublin's population is 1,000,000. Dublin has 1,000,000 inhabitants. One million people live in Dublin. [Dublin] Its population is one million. La población de Dublín es un millón.





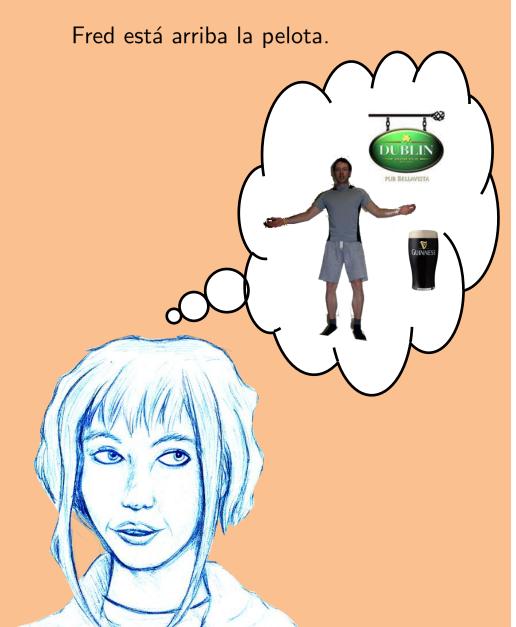
MULTIPLE MEANINGS FOR THE SAME SAYING ...

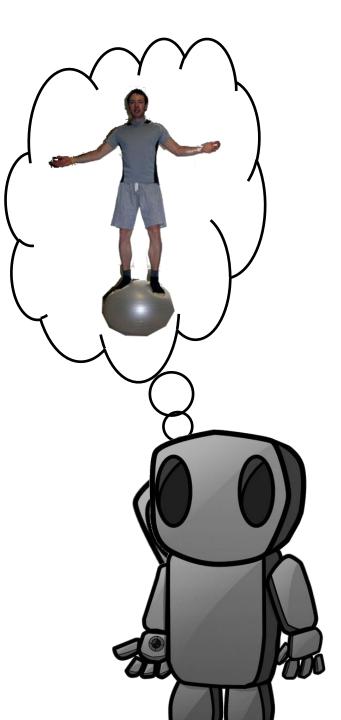
Sherlock saw the man using binoculars.





NOT SAYING WHAT IS MEANT ...



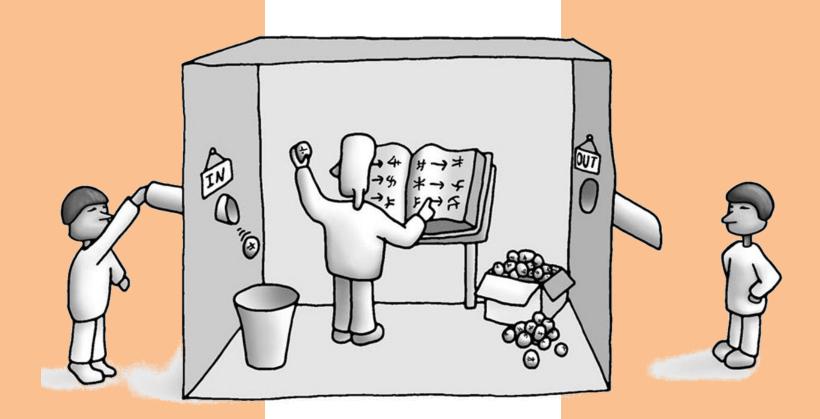


The Semantic Gap





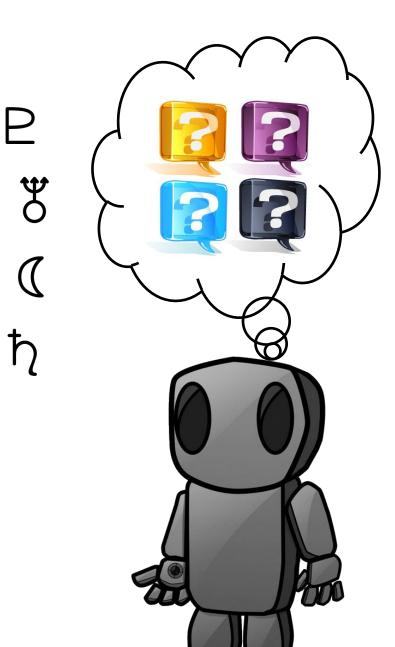
WHAT IF WE COULD "STRUCTURE" EVERYTHING ...

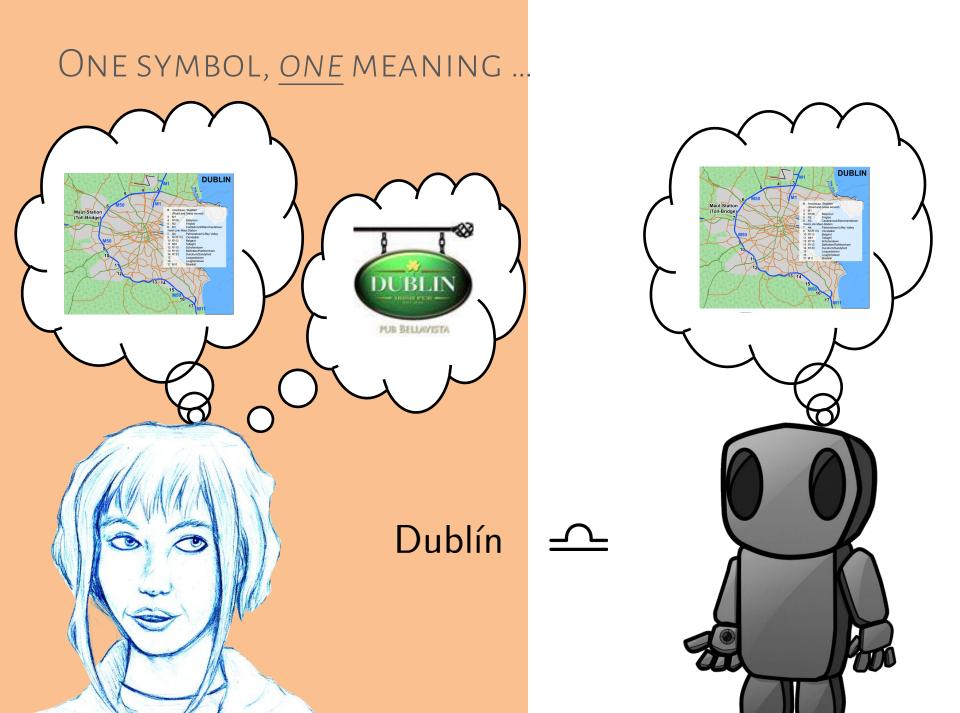


ONE SYMBOL, ONE MEANING ...

IE







ONE (SIMPLE) WAY TO SAY ONE THING ...

Dublin's population is one million. Dublin has a population of one million. Dublin's population is 1,000,000. Dublin has 1,000,000 inhabitants. One million people live in Dublin. [Dublin] Its population is one million. La población de Dublín es un millón. (Dublin,population,1000000)

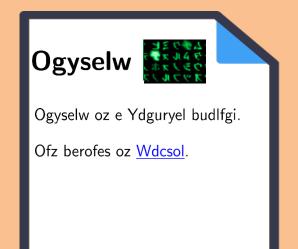


(^w₀, o', 1000000)



SEARLE'S CHINESE ROOM (NATURAL LANGUAGE)

http://ex.org/Ogyselw



http://ex.org/Wdcsol



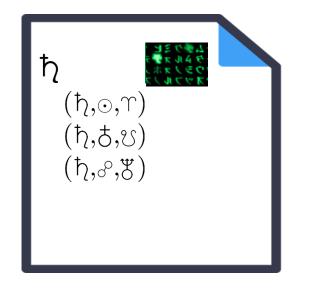
INPUT: "Nhef oz fhy rurdsefoul up fhy berofes up Ogyselw?"

... what should the output be?

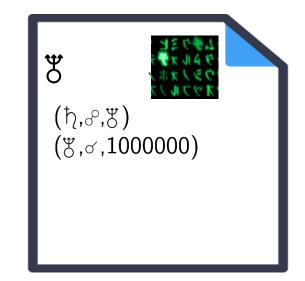
 $OUTPUT: \ ``uly \ mossoul''$

SEARLE'S CHINESE ROOM (SYMBOLIC)

http://ex.org/ħ



http://ex.org/ ∇



INPUT: " $(\uparrow, \circ, x), (x, \sigma, y)$?"

... what should the output be?

OUTPUT: $\{(x \mapsto \mathcal{Z}, y \mapsto 1000000)\}$



SEARLE'S CHINESE ROOM (SYMBOLIC)

http://ex.org/Ireland



http://ex.org/Dublin



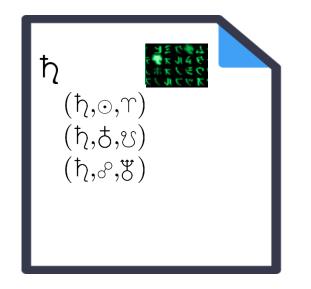
INPUT: "(Ireland, capital, x), (x, population, y)?"

... what should the output be?

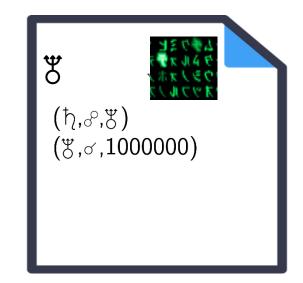
OUTPUT: $\{(x \mapsto \mathsf{Dublin}, y \mapsto 100000)\}$



http://ex.org/ħ



http://ex.org/∛



INPUT: " (x, \odot, y) ?"

... what should the output be?

Output: $\{(x \mapsto h, y \mapsto \uparrow)\}$



SEARLE'S CHINESE ROOM II

http://ex.org/Ireland



http://ex.org/Dublin



INPUT: "(*x*,partOf,*y*)?"

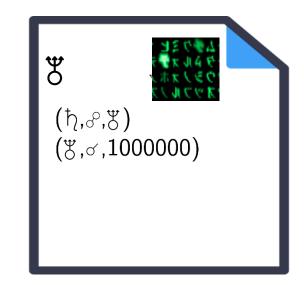
... what should the output be?

OUTPUT: { $(x \mapsto \text{Ireland}, y \mapsto \text{Europe})$ } $(x \mapsto \text{Dublin}, y \mapsto \text{Ireland})$?



http://ex.org/ħ

わ (れ,⊙,↑) (れ,ま,v) (れ,♂,♂) http://ex.org/~



Rule: " $(b, \circ, a) \rightarrow (a, \circ, b)$ "

INPUT: " (x, \odot, y) ?"

... what should the output be?

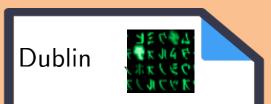
OUTPUT: $\{(x \mapsto h, y \mapsto f), (x \mapsto g, y \mapsto h)\}$



http://ex.org/Ireland



http://ex.org/Dublin



(Ireland,capital,Dublin) (Dublin,population,1000000)

RULE: " $(b, capital, a) \rightarrow (a, partOf, b)$ "

INPUT: "(*x*,partOf,*y*)?"

... what should the output be?

OUTPUT: { $(x \mapsto \text{Ireland}, y \mapsto \text{Europe}),$ $(x \mapsto \text{Dublin}, y \mapsto \text{Ireland})$ } $(x \mapsto \text{Dublin}, y \mapsto \text{Europe})?$



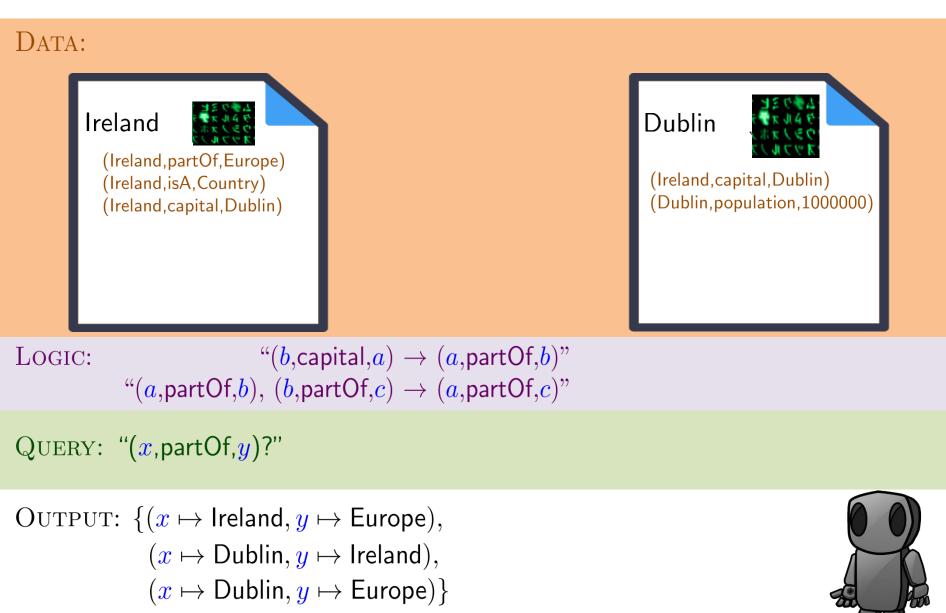
http://ex.org/ħ Y ħ (\hbar, \odot, \uparrow) (ኪ,ぷ,쀻) (れ,ま,ど) (g, , 100000) (ħ,₀°,\) $``(b,\circ,a) ightarrow (a,\circ,b)"$ RULES: $``(a,\odot,b), (b,\odot,c) \rightarrow (a,\odot,c)"$ INPUT: " (x, \odot, y) ?" ... what should the output be? OUTPUT: $\{(x \mapsto h, y \mapsto \uparrow), (x \mapsto F, y \mapsto h), (x \mapsto F, y \mapsto \uparrow)\}$

http://ex.org/∛

SEARLE'S CHINESE ROOM II

http://ex.org/Ireland http://ex.org/Dublin Ireland Dublin (Ireland, partOf, Europe) (Ireland, capital, Dublin) (Ireland, isA, Country) (Dublin, population, 1000000) (Ireland, capital, Dublin) " $(b, capital, a) \rightarrow (a, partOf, b)$ " RULES: "(a,partOf,b), (b,partOf,c) \rightarrow (a,partOf,c)" INPUT: "(x, partOf, y)?" ... what should the output be? OUTPUT: {($x \mapsto$ Ireland, $y \mapsto$ Europe), $(x \mapsto \mathsf{Dublin}, y \mapsto \mathsf{Ireland}),$ $(x \mapsto \mathsf{Dublin}, y \mapsto \mathsf{Europe})$

SEMANTIC WEB: DATA, LOGIC, QUERY



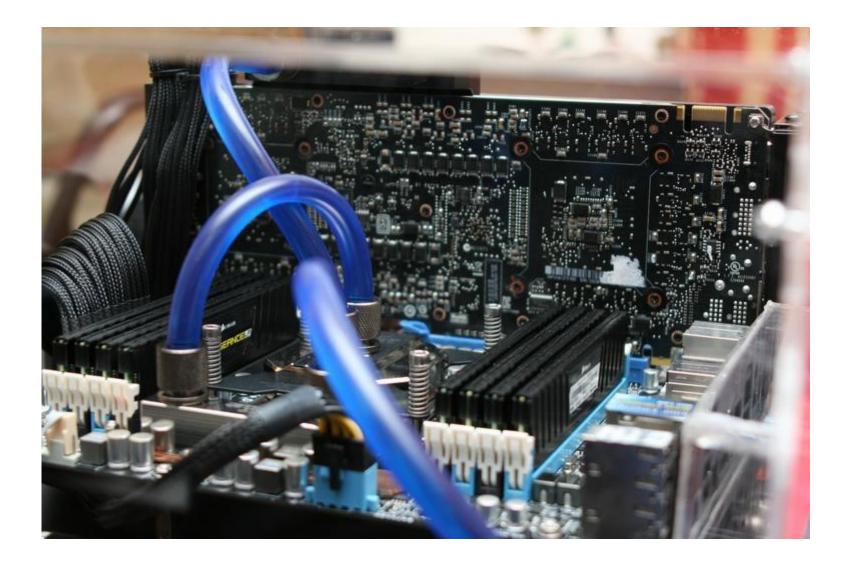
THE SEMANTIC WEB NOW?

THE SEMANTIC WEB IS NOW ABOUT 20 YEARS OLD



... so where is it then?

HIDDEN WITHIN THE WEB ...



WIKIDATA: A WIKIPEDIA FOR DATA

What is Wikidata?



Main page Community portal Project chat Create a new Item Recent changes Random Item Query Service Nearby Help

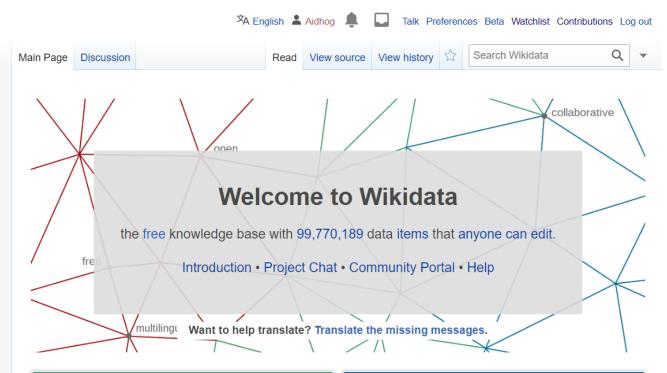
Donate

Lexicographical data

Create a new Lexeme Recent changes Random Lexeme

Tools

What links here Related changes Special pages Permanent link Page information Wikidata item



Welcome!

Wikidata is a free and open knowledge base that can be read and edited by both humans and machines.

Wikidata acts as central storage for the **structured data** of its Wikimedia sister projects including Wikipedia, Wikivoyage, Wiktionary, Wikisource, and others.

Learn about data

New to the wonderful world of data? Develop and improve your data literacy through content designed to get you up to speed and feeling comfortable with the fundamentals in no time.



WHY IS WIKIDATA?



Main page Community portal Project chat Create a new Item Recent changes Random Item Query Service Nearby Help

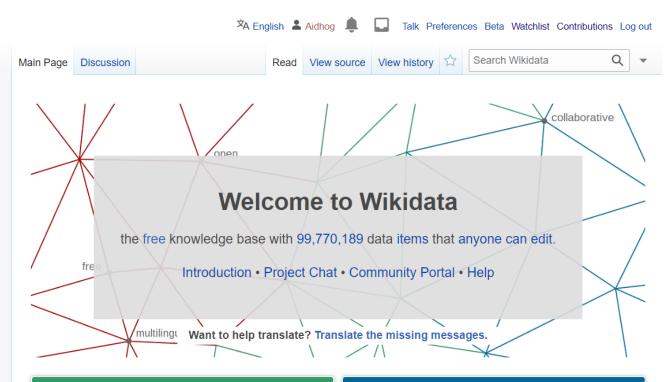
Donate

Lexicographical data

Create a new Lexeme Recent changes Random Lexeme

Tools

What links here Related changes Special pages Permanent link Page information Wikidata item



Welcome!

Wikidata is a free and open knowledge base that can be read and edited by both humans and machines.

Wikidata acts as central storage for the **structured data** of its Wikimedia sister projects including Wikipedia, Wikivoyage, Wiktionary, Wikisource, and others.

Learn about data

New to the wonderful world of data? Develop and improve your data literacy through content designed to get you up to speed and feeling comfortable with the fundamentals in no time.



PROBLEM 1: DIFFERENT LANGUAGE VERSIONS MANUALLY EDITED BY USERS

Alexis Sánchez



Sánchez with Chile in 2013

Personal information				
Full name	Alexis Alejandro Sánchez Sánchez ^[1]			
Date of birth	19 December 1988 (age 26) ^[1]			
Place of birth	Tocopilla, Chile ^{[3][2]}			

	National team [‡]		
2007	Chile U20	12	(2)
2006–	Chile	82	(26)



Nombre completo Alexis Alejandro Sánchez Sánchez



Alexis Sánchez im Dezember 2011

Spielerinformationen

Voller Name Alexis Alejandro Sánchez Sánchez Geburtstag 19. Dezember 1988 Geburtsort Tocopilla, Chile

Carrera internacional					
Selección	Chile				
Part. (goles)	82 (26)				
Debut	2006				

Nationalmannschaft ²					
2007	Chile U-20				
2006-	Chile	76 (25)			

PROBLEM 2: COMPLEX LISTS OF THINGS MANUALLY EDITED BY USERS



Article Talk

Solution Contributions Create account Log in

Read	Edit	View history	Search	Q

Chile national football team

From Wikipedia, the free encyclopedia

at club level.

Most capped players [edit]

As of September 1, 2016 Players in **bold** are still active, at least

Top goalscorers [edit]

As of September 1, 2016

Players in **bold** are still active, at least at club level.

#	Name	International Career	Caps	Goa	ıls#	Name	International Career	Goals	Caps
1.	Claudio Bravo	2004 –	106	0	1.	Marcelo Salas	1994–2007	37	70
2.	Alexis Sánchez	2006 –	102	34	ŀ	lván Zamorano	1987–2001	34	69
3.	Gary Medel	2007 –	96	7	2.	Alexis			
4.	Gonzalo Jara	2006 –	95	3		Sánchez (list)	2006 –	34	102

ALEXIS SCORES A GOAL ...



Now an army of human editors has to manually update a bunch of articles: different languages, lists, ...

Solution: Wikidata



Main page Community portal Project chat Create a new Item Recent changes Random Item Query Service Nearby Help

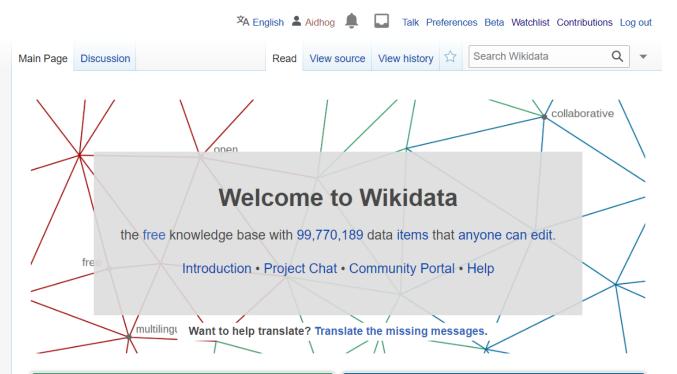
Donate

Lexicographical data

Create a new Lexeme Recent changes Random Lexeme

Tools

What links here Related changes Special pages Permanent link Page information Wikidata item



Welcome!

Wikidata is a free and open knowledge base that can be read and edited by both humans and machines.

Wikidata acts as central storage for the **structured data** of its Wikimedia sister projects including Wikipedia, Wikivoyage, Wiktionary, Wikisource, and others.

Learn about data

New to the wonderful world of data? Develop improve your data literacy through content de get you up to speed and feeling comfortable fundamentals in no time.



Main page

Project chat

Item by title

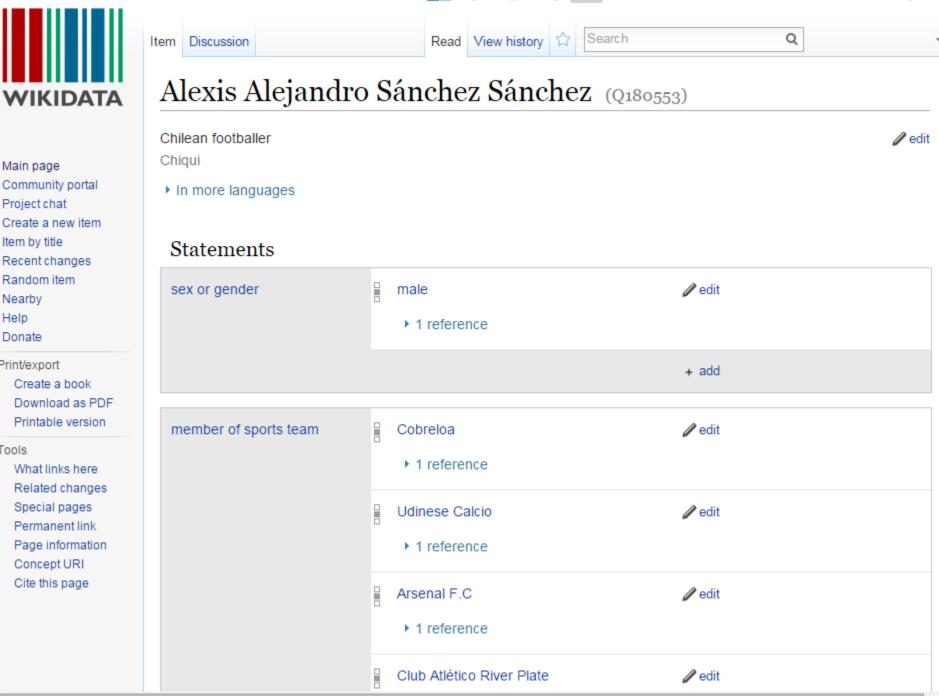
Nearby Help

Donate

Tools

- I

Print/export



A あ English S Aidhog 🔍 0

Talk Preferences Beta Watchlist Contributions Log out



Página principal	
Portal de la comunidad	
Café	

Crear un elemento nuevo

Elemento por título

Cambios recientes

Elemento aleatorio

Cercanos

Ayuda

Donaciones

Imprimir/exportar

Crear un libro Descargar como PDF

Versión para imprimi

Herramientas

- II

Lo que enlaza aquí Cambios relacionados Páginas especiales Enlace permanente Información de la página Concept URI Citar esta página

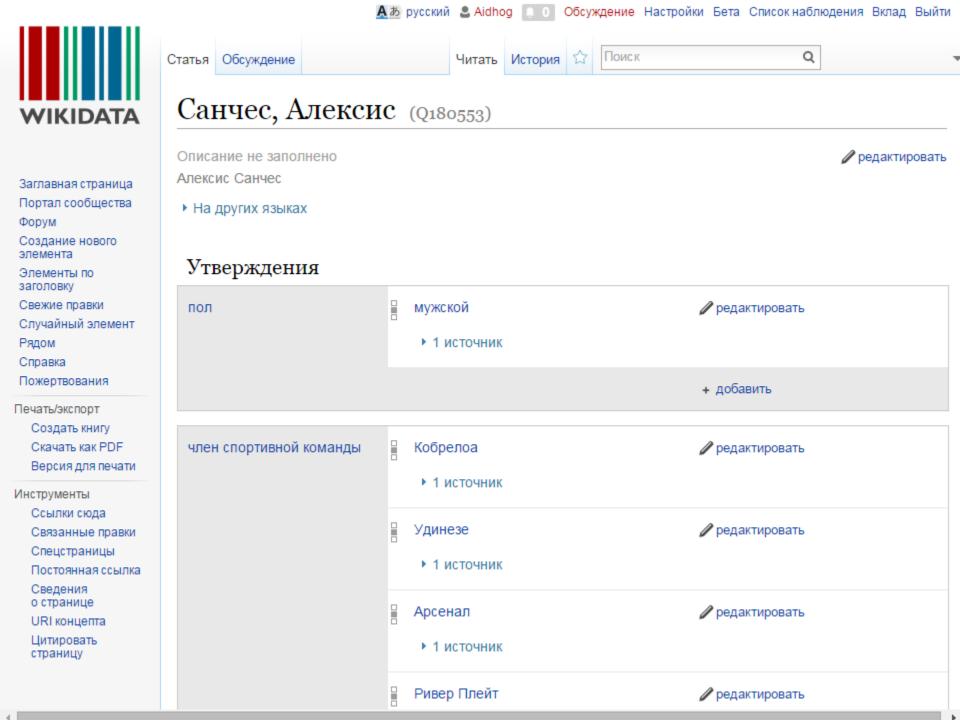
		A b es	pañol 💄 Aidhog	🌲 0 Disc	usión	Preferencias	Beta Listado	e seguimiento	Contribuciones	Salir
Elemento	Discusión		Leer	Ver historial	☆	Buscar		Q		
Alex	Alexis Sánchez (Q180553)									
	a chileno	nchez I AS9 I /	Alexis Alejandro	Sánchez S	ánche		levis Aleiandı	ro Sánchez I	2	editar
		anchez Alexis	-	Sanchez Sa	ancin			o Ganeriez j	AICAIS	
► En m	ás idiomas									
Decl	laracion	es								
sexo			masculino				🥒 editar			
	► 1 referencia									
							+ añadir	r		

miembro del equipo deportivo	Club de Deportes Cobreloa 1 referencia	🖉 editar
	Udinese Calcio 1 referencia	🥒 editar
	Arsenal Football Club 1 referencia 	🥒 editar



- II

	Ab	Deutsch 💄 Aidhog 🏮 0 Diski	ussion Einstellungen Beta Beobachtun	igsliste Beiträge Abmelden
WIKIDATA	Datenobjekt Diskussion Alexis Sánchez (Lesen Versionsgeschichte (Q180553)	☆ Suchen	Q
Hauptseite Gemeinschaftsportal Forum Ein neues Datenobjekt erstellen Datenobjekte nach Titel	chilenischer Fußballspieler Alexis Alejandro Sánchez Sánche In weiteren Sprachen Aussagen	iez Alexis Sanchez Alexis A	lejandro Sanchez Sanchez	🖉 bearbeiten
Letzte Änderungen Zufälliges Datenobjekt In der Nähe Hilfe Spenden	Geschlecht	männlich1 Fundstelle	🖉 bearbeiten	
Drucken/exportieren			+ hinzufügen	
Buch erstellen Als PDF herunterladen Druckversion Werkzeuge Links auf diese Seite Änderungen an verlinkten Seiten Spezialseiten Permanenter Link	Mitglied von Sportmannschaft oder -verein	CD Cobreloa	🖉 bearbeiten	
		Udinese Calcio 1 Fundstelle 	/ bearbeiten	
Seiteninformationen Konzept-URI Seite zitieren		FC Arsenal 1 Fundstelle 	🥒 bearbeiten	
		CA River Plate	🖉 bearbeiten	





പ്രധാന താൾ സാമൂഹികകവാടം Project chat പുതിയൊരു ഇനം സ്പ്ഷ്പിക്കുക് ഇനം തലക്കട്ടനുസരിച്ച് സമീപകാല മാറ്റങ്ങൾ എതെങ്കിലും താൾ സമീപസ്ഥം സഹായം സംഭാവന അച്ചടിയ്ക്കുക/ കയ്റ്റുമതി ചെയ്യുക പുസ്തകം സ്പ്ഷ്പിക്കുക PDF ആയി

PDF ആയി ഡൗൺലോഡ് ചെയ്യുക അച്ചടിരൂപം

ഉപകരണങ്ങൾ **സെ**

ഈ താളിലേക്കുള്ള കണ്ണികൾ അനുബന്ധ മാറ്റങ്ങൾ പ്രത്യേക താളുകൾ

സ്ഥിരംകണ്ണി താഭിന്റെ

അലക്സിസ് സാഞ്ചസ് (Q180553)

വായിക്കുക നാൾവഴി കാണുക 🟠

വിവരണമൊന്നും നിർവചിച്ചിട്ടില്ല അപരനാമങ്ങളൊന്നും കണ്ടെത്താനായില്ല.

കൂടുതൽ ഭാഷകളിൽ

Statements

ലേഖനം സംവാദം

ലിംഗം		പുരുഷൻ ▶ 1 സ്രോതസ്സ്	🖉 തിരുത്തുക
			+ ചേർക്കുക
member of sports team <i>ഇംഗ്ലീഷ്</i>		Cobreloa <i>ഇംഗ്ലീഷ്</i> ▶ 1 സ്രോതസ്സ്	🖋 തിരുത്തുക
		Udinese Calcio <i>ഇംഗ്ലീഷ്</i> ▶ 1 സ്രോതസ്സ്	🖉 തിരുത്തുക
		ആഴ്സണൽ എഫ്.സി. ▶ 1 സ്രോതസ്സ്	🖉 തിരുത്തുക
		Club Atlético River Plate <i>உலி</i> ഷ്	🖉 തിരാത്താക

തിരയുക

🥒 തിരുത്തുക

Q

What does Wikidata describe?

(71,611,020)

human: 6,376,879 (8.9%) taxon: 2,726,046 (3.8%) administrative territorial entity: 1,943,285 (2.7%) architectural structure: 3,159,472 (4.4%) occurrence: 3,898,674 (5.4%) chemical compound: 1,188,724 (1.7%) film: 294,370 (0.4%) thoroughfare: 630,794 (0.9%) astronomical object: 4,601,733 (6.4%) Wikimedia list article: 404,454 (0.6%) Wikimedia disambiguation page: 1,358,230 (1.9%) Wikinews article: 195,900 (0.3%) scholarly article: 22,574,314 (31.5%) other P31/P279: 18,284,676 (25.5%) no P31/P279: 3,973,469 (5.5%)

• 22,986 active users

https://www.wikidata.org/wiki/Wikidata:Statistics

Use-case: Info-Boxes



The Free Encyclopedia

Main page Contents Featured content Current events Random article Donate to Wikipedia Wikipedia store

Interaction

Help About Wikipedia Community portal Recent changes Contact page Article Talk

Atacama Pathfinder Experiment

From Wikipedia, the free encyclopedia

The Atacama Pathfinder Experiment (APEX) is a radio telescope 5,100 meters above sea level, at the Llano de Chajnantor Observatory in the Atacama desert in northern Chile, 50 km east of San Pedro de Atacama built and operated by 3 European research institutes. The main dish has a diameter of 12 m and consists of 264 aluminium panels with an average surface accuracy of 17 micrometres (rms). The telescope was officially inaugurated on September 25, 2005.

The APEX telescope is a modified ALMA (Atacama Large Millimeter Array) prototype antenna and is at the site of the ALMA observatory. APEX is designed to work at sub-millimetre wavelengths, in the 0.2 to 1.5 mm range — between infrared light and radio waves — and to find targets that ALMA will be able to study in greater detail. Submillimetre astronomy provides a window into the cold, dusty and distant Universe, but the faint signals from space are heavily absorbed by water vapour in the Earth's atmosphere. Chajnantor was chosen as the location for such a telescope because the region is one of the driest on the planet and is more than 750

Atacama Pathfinder Experiment



USE-CASE: INFO-BOXES



Contents Featured content Current events

Main page Community portal Project chat Create a new item Recent changes Random item Query Service Nearby Help Donate

Tools

What links here Related changes Special pages Article Talk

Item Discussion

Atacama Pathfinder Experiment

From Wikipedia, the free encyclopedia

The Atacama Pathfinder Experiment (APEX) is a radio telescope 5,100 meters above sea level, at the Llano de Chajnantor Observatory in the Atacama desert in northern Chile, 50 km east of San Pedro de Atacama built and operated by 3 European research institutes. The main dish has a diameter of 12 m and consists of 264 aluminium panels with an average surface accuracy of 17 micrometres (rms). The telescope was officially

Search Wikidata

Q

Atacama Pathfinder Experiment (Q753076)

Read View history



Atacama Pathfinder Experiment

The APEX telescope					
Observatory	Llano de Chajnantor Observatory 🖉				
Location(s)	Atacama Desert, Chile 🖉				
Coordinates	Q 23°00′21″S 67°45′33″W				
Organization	Edit this at Wikidata Observatory Max Planck Institute for Radio Astronomy Onsala Space Observatory				
Altitude	5,100 m (16,700 ft) 🖉				
Wavelength	0.2, 1.5 mm (1.50, 0.20 THz)				
First light	2004 🖉				
Telescope style	Cassegrain reflector Cosmic microwave background experiment Radio telescope 🖉				

USE-CASE: QUALITY CHECKS

List of all person who do not have an age between 0 and 130. Update: 22:21, 10 September 2018 (UTC)

Item	♦ Birth ♦	Death 🔶	Age 🗢
Tuti Yusupova (Q1038827)	1880-07-01	2015-03-28	134
Karni Mata (Q1106783)	1387-10-02	1538-03-23	150
no label (Q11556831)	1185-01-01	1392-01-01	207
Minamoto no Chikayuki (Q11564306)	1185-01-01	1333-01-01	148
Bir Narayan Chaudhuri (Q11854281)	1857-00-00	1998-04-20	141
no label (Q12123094)	1091-00-00	1228-00-00	137
no label (Q12218744)	1819-00-00	1954-00-00	135
Maftei Pop (Q12734691)	1804-01-01	1952-03-15	148
Habib Miyan (Q1365575)	1868-05-20	2008-08-19	140
Egyō (Q1392070)	805-01-01	1185-01-01	380
Xu Xun (Q1428729)	239-01-01	374-01-01	135
Zaro Aga (Q148028)	1777-01-01	1934-06-29	157
Charles Étienne Guillaume Blandin de Chalain (Q15967550)	1740-06-07	1958-01-01	217
Javier Pereira (Q15999178)	1789-00-00	1958-00-00	169
Chen Jun (Q16077971)	881-00-00	1324-00-00	443
Jules Granier (Q16842647)	1770-01-01	1906-04-07	136
Genson (Q18115051)	1700-01-01	1950-01-01	250
Liutwin (Q18222784)	1200-01-01	1350-01-01	150
Salah (Q1827950)	-2068-00-00	-1635-00-00	433

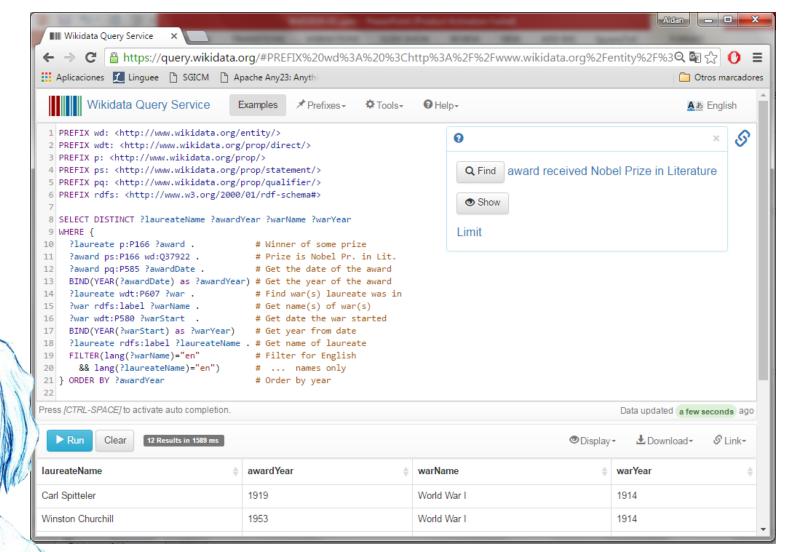
USE-CASE: QUALITY CHECKS

List of all person who do not have an age between 0 and 130. Update: 22:21, 10 September 2018 (UTC)

Item	♦ Birth ♦	Death 🔶	Age 🗢
Tuti Yusupova (Q1038827)	1880-07-01	2015-03-28	134
Karni Mata (Q1106783)	1387-10-02	1538-03-23	150
no label (Q11556831)	1185-01-01	1392-01-01	207
Minamoto no Chikayuki (Q11564306)	1185-01-01	1333-01-01	148
Julius Fessler (Q55677113)	1982-05-04	1937-10-27	-45
Max Hallbauer (Q55678777)	1851-12-04	1818-10-08	-34
Ernst Keiter (Q55679293)	1943-10-28	1907-10-30	-36
Franz Friedrich Theodor Steinhauer (Q55680095)	1849-12-25	1822-12-21	-28
Johann Joseph Schoder (Q55680781)	1918-12-07	1884-12-12	-34
Paul Kaspar Helbling (Q55681272)	2012-06-05	2011-06-20	-1
Gustaf Bolinger (Q55683866)	1988-12-19	1957-07-16	-32
Joseph Marcus Jaffé (Q55683961)	1867-03-11	1841-04-12	-26
Johann Baptist Hau (Q55683983)	1782-08-14	1758-12-26	-24
Wilhelm Klaubert (Q55684184)	1557-08-17	1526-10-03	-31
Dieterich Johann Krüger (Q55684826)	1742-06-23	1726-08-22	-16
Alfred Reichenbecher (Q55684894)	1884-01-26	1664-03-05	-220
Johann Ludwig Winckler (Q55901640)	1963-05-08	1767-08-08	-196
Johann Heremberck (Q55902890)	1811-01-01	1489-01-01	-322
Salah (Q1827950)	-2068-00-00	-1635-00-00	433

USE-CASE: QUERY SERVICE





USE-CASE: QUERY SERVICE



Aidan 🗖 🗖 🗙 IIII Wikidata Query Service 🛛 🗙 🚺 🔒 https://query.wikidata.org/#PREFIX%20wd%3A%20%3Chttp%3A%2F%2Fwww.wikidata.org%2Fentity%2F%3Q 🗟 🛠 🕧 🚍 C ← 👖 Aplicaciones 🌠 Linguee 🌓 SGICM 🌓 Apache Any23: Anythi Otros marcadores rwan พนิเวิครอช rwanstant . # dec date the war started 17 BIND(YEAR(?warStart) as ?warYear) # Get year from date ?laureate rdfs:label ?laureateName . # Get name of laureate 18 FILTER(lang(?warName)="en" # Filter for English 19 20 && lang(?laureateName)="en") # ... names only 21 } ORDER BY ?awardYear # Order by year 22 Press [CTRL-SPACE] to activate auto completion. Data updated a few seconds ago Run Clear 12 Results in 1589 ms Oisplay • L Download -S Link+ laureateName awardYear warName warYear 1919 World War I Carl Spitteler 1914 Winston Churchill 1953 World War I 1914 1954 World War I Ernest Hemingway 1914 Ernest Hemingway 1954 World War II 1939 Jean-Paul Sartre 1964 Algerian War 1954 Jean-Paul Sartre 1964 World War II 1939 Heinrich Böll 1972 World War II 1939 Eugenio Montale 1975 World War I 1914 William Golding 1983 World War II 1939 Claude Simon 1985 Spanish Civil War 1936 Camilo José Cela 1989 Spanish Civil War 1936 Günter Grass 1999 World War II 1939



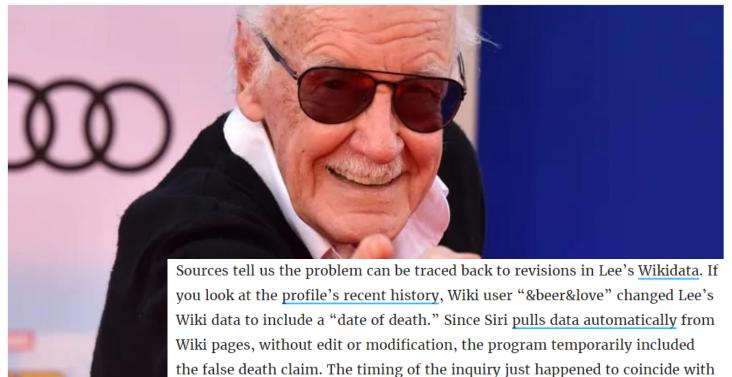
USED IN APPLICATIONS LIKE SIRI ...

Siri Erroneously Told People Stan Lee Was Dead



Beth Elderkin 7/03/18 2:45pm • Filed to: STAN LEE ~





the false information being present at that point.

Stan Lee at the premiere of *Spider-Man: Homecoming*. Photo: Alberto E. Rodriguez (Getty Images)

Google's Knowledge Graph

Google's Knowledge Panel

https://www.google.cl/?gfe_rd=cr&ei=AfHMV9-YCMgnxgT350 C

👖 Aplicaciones 🌠 Linguee 🎦 SGICM 🎦 Apache Any23: Anyth

Google sully prudhomme

G sully prudhomme - Goog X

Search tools Images Videos Books News More -

About 372,000 results (0.32 seconds)

Sully Prudhomme - Wikipedia, the free encyclopedia https://en.wikipedia.org/wiki/Sully Prudhomme *

René François Armand (Sully) Prudhomme was a French poet and essayist. He was the first ever winner of the Nobel Prize in Literature in 1901. Born in Paris ... Early life · Writing · Nobel Prize · Death

Sully Prudhomme - Wikipedia, la enciclopedia libre

https://es.wikipedia.org/wiki/Sully Prudhomme * Translate this page

René François Armand (Sully) Prudhomme también conocido como Sully Prudhomme (París, Francia, 16 de marzo de 1839 - Châtenay-Malabry, Francia, 6 de ...

Sully Prudhomme - Biographical - Nobelprize.org

www.nobelprize.org/nobel_prizes/literature/laureates/1901/prudhomme-bio.html * Rene Francois Armand Prudhomme (1839-1907) was the son of a French shopkeeper. ... Sully

Prudhomme was a member of the «Conference La Bruyère», ...

Sully Prudhomme - Nobelprize.org

https://www.nobelprize.org/nobel_prizes/literature/laureates/1901/ *

The Nobel Prize in Literature 1901 was awarded to Sully Prudhomme 'in special recognition of his poetic composition, which gives evidence of lofty idealism, ...

Sully Prudhomme | French poet | Britannica.com

https://www.britannica.com/biography/Sully-Prudhomme *

Sully Prudhomme, pseudonym of René-François-Armand Prudhomme (born March 16, 1839, Parisdied Sept. 7, 1907, Châtenay, France) French poet who ...

Sully Prudhomme - NNDB.com

www.nndb.com/people/297/000098003/ *

It was at this moment that the small circle of which Leconte de Lisle was the center were preparing the Parnasse, to which Sully Prudhomme contributed several

Cully Drudhamma I Definition of Cully Drudhamma by Mamiam Wahater



Sully Prudhomme

Poet

René François Armand Prudhomme was a French poet and essayist. He was the first ever winner of the Nobel Prize in Literature in 1901. Wikipedia

Born: March 16, 1839, Paris, France

Died: September 6, 1907, Châtenay-Malabry, France

Books: Les vaines tendresses

Awards: Nobel Prize in Literature

Theodor

People also search for







- -

Aidan



View 10+ more

Leconte de Lisle

Mommsen

Paul Verlaine Gabriel Fauré

Frédéric Mistral

USING SEMANTIC WEB KNOWLEDGE-BASES

From Freebase to Wikidata: The Great Migration

Thomas Pellissier Tanon Google, San Francisco, USA thomas@pellissier-tanon.fr vrandecic@google.com

Denny Vrandečić Google, San Francisco, USA

Sebastian Schaffert Google, Zürich, Switzerland schaffert@google.com

Thomas Steiner Google, Hamburg, Germany tomac@google.com

Lydia Pintscher Wikimedia, Berlin, Germany lydia@pintscher.de

ABSTRACT

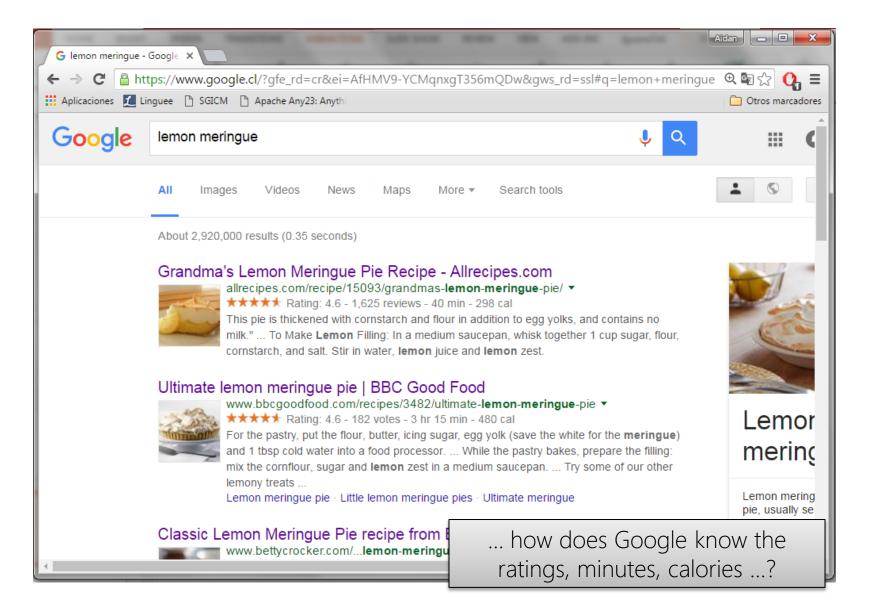
Collaborative knowledge bases that make their data freely available in a machine-readable form are central for the data strategy of many projects and organizations. The two major collaborative knowledge bases are Wikimedia's Wikidata and Google's Freebase. Due to the success of Wikidata, Google decided in 2014 to offer the content of Freebase to the Wikidata community. In this paper, we report on the ongoing transfer efforts and data mapping challenges, and provide an analysis of the effort so far. We describe the Primary Sources Tool, which aims to facilitate this and future data migrations. Throughout the migration, we have gained deep insights into both Wikidata and Freebase, and share and discuss detailed statistics on both knowledge bases.

One such collaborative knowledge base is Freebase, publicly launched by Metaweb in 2007 and acquired by Google in 2010. Another example is Wikidata, a collaborative knowledge base developed by Wikimedia Deutschland since 2012 and operated by the Wikimedia Foundation. Due to the success of Wikidata, Google announced in 2014 their intent to shut down Freebase and help the community with the transfer of Freebase content to Wikidata [10].

Moving data between two knowledge bases that do not share a similar design is usually a problematic task and requires the careful mapping between their structures. The migration from Freebase to Wikidata was no exception to this rule: we encountered a number of to-be-expected structural challenges. However, even more demanding was the cultural difference between the two involved communities.

Google's Rich Snippets

FANCY-LOOKING SEARCH RESULTS ...



PUBLISHERS ADD STRUCTURED DATA!

```
itemtype="http://schema.org/NutritionInformation">

<span class="nutrition_label">kcal</span>
<span class="nutrition_value" itemprop="calories">480</span>
[...]
```

✓ Publishers get more clicks on their results

✓ Google gets data to make fancy results

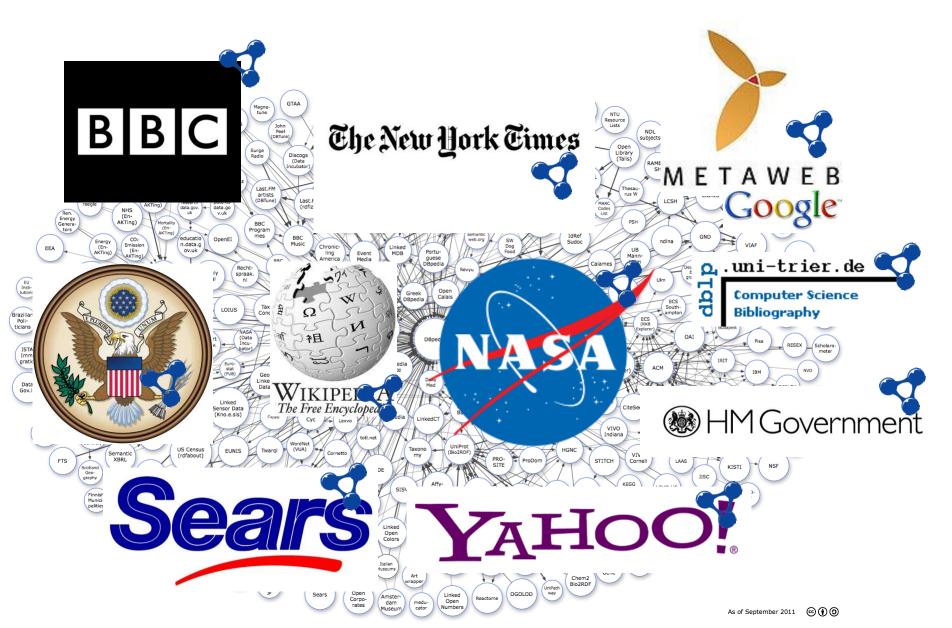
Standards for Open Data: Linked Open Data

HOW TO PUBLISH OPEN DATA?



5-Star Linking Open Data Scheme

The Linked Data Cloud



About the course ...

TOPICS COVERED

- RDF (triple-based data model)
- RDFS/OWL (ontological languages)
- SPARQL (query language)
- Linked Data / Web of Data
- RDB2RDF (importing databases to Sem. Web.)
- Shapes (validating RDF data)

Research course



STRUCTURE OF THE COURSE

- Each week:
 - Class on Monday (learn concepts)
 - Lab on Wednesday (see concepts in practice)
 - Auxiliar session on Friday (Q&A)
- Marking structure:
 - 70% labs
 - 20% project
 - 10% reading group

BIBLIOGRAPHY

Aidan Hogan The Web of Data

This book concisely brings together the key standards and best practices relating to modelling, querying, validating and linking machine-readable data and semantics on the Web. Alongside practical examples and formal definitions, the book shows how these standards contribute to – and have been used thus far on – the "Web of Data": a machine readable evolution of the Web marked by increased automation, enabling powerful Web applications capable of discovering, cross-referencing, and organising data from numerous websites in a matter of seconds.

The book is divided into nine chapters, the first of which highlights the fundamental shortcomings of the current Web that illustrate the need for increased machine readability. The next chapter outlines the core concepts of the "Web of Data", discussing use-cases on the Web where they have already been deployed. "Resource Description Framework (RDF)" describes the graph-structured data model proposed by the Semantic Web community as a common data model for the Web. The chapter on "RDF Schema (RDFS) and Semantics" presents a lightweight ontology language used to define an initial semantics for RDF graphs. In turn, the chapter "Web Ontology Language (OWL)" elaborates on a much more expressive ontology language built upon RDFS. In "SPARQL Query Language" a language for querying and updating RDF graphs is described. "Shape Constraints and Expressions (SHACL/ShEx)" introduces two languages for describing the expected structure of - and expressing constraints over - RDF graphs for the purposes of validation. "Linked Data" discusses the principles and best practices by which interlinked (RDF) data can be published on the Web, and how they have been adopted. The final chapter highlights open problems and concludes with a general discussion on the future of the Web of Data.

The book is intended for students, researchers and advanced practitioners interested in learning more about the Web of Data, and about closely related topics such as the Semantic Web, Knowledge Graphs, Linked Data, Graph Databases, Ontologies, etc. Offering a range of accessible examples and exercises, it can be used as a textbook for students and other newcomers to the field. It can also serve as a reference handbook for researchers and developers, as it offers up-to-date details on key standards (RDF, RDFS, OWL, SPARQL, SHACL, ShEx, RDB2RDF, LDP), along with formal definitions and references to further literature. The associated website webofdatabook.org offers a wealth of complementary material, including solutions to the exercises, slides for classes, interactive examples, and a section for comments and questions. Aidan Hogar

ً

The Web of Data

Aidan Hogan

The Web of Data

Outcomes: Learn About the Semantic Web!

- An ongoing research topic here in the DCC
- Apply database, logic, Al, etc., to the Web
- Mix of theory and practical exercises
- The future of the Web?

QUESTIONS?

