# CC7220-1
# La Web de Datos
# Primavera 2020

## Lecture 1: Introduction

Aidan Hogan
aidhog@gmail.com

# The Web

# 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?


"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."

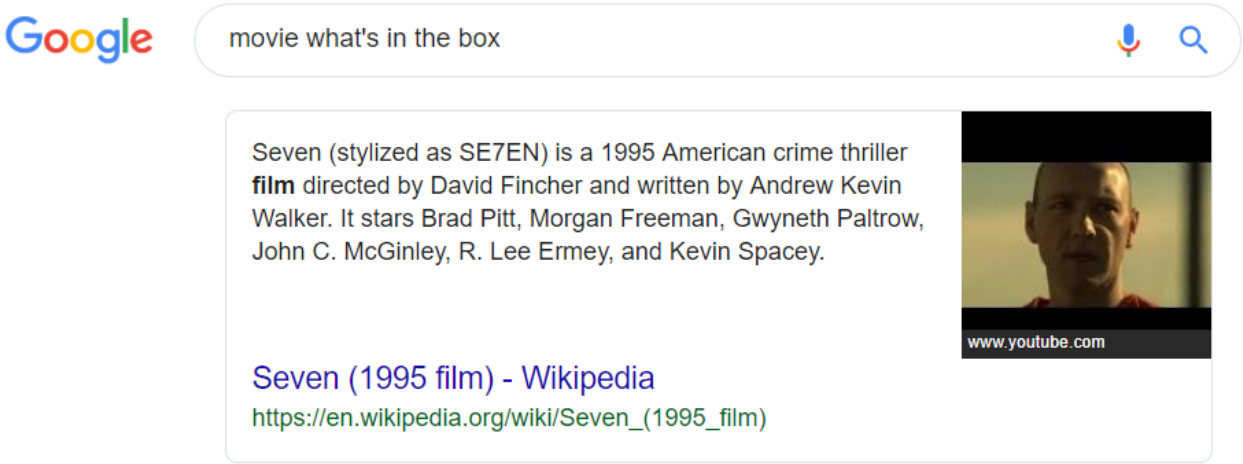
*– Berners-Lee et al. (2001) "The S[illegible]"*
*Sci. American, 2[illegible]3.*

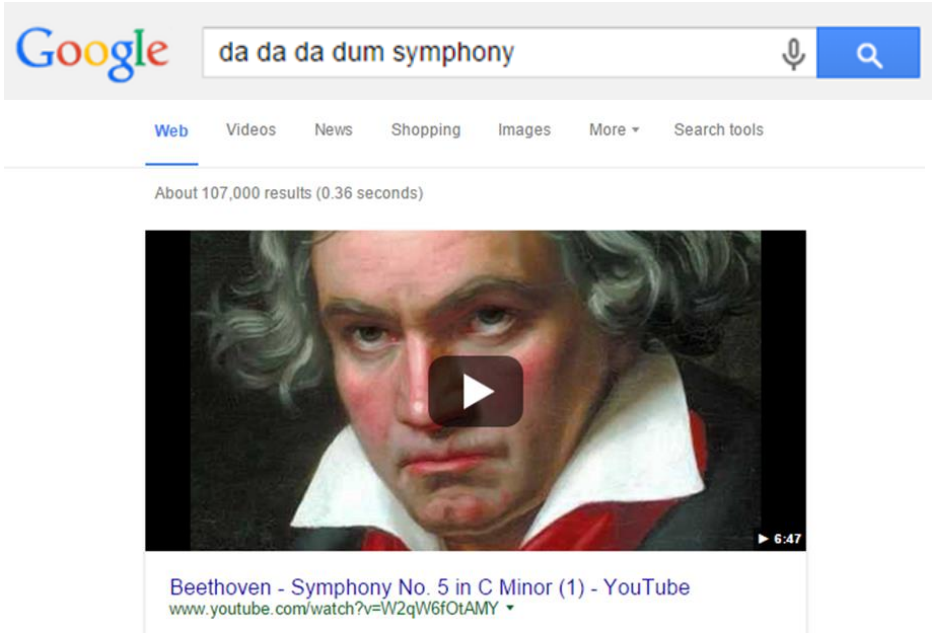
Consider answering: "What is the Web?"

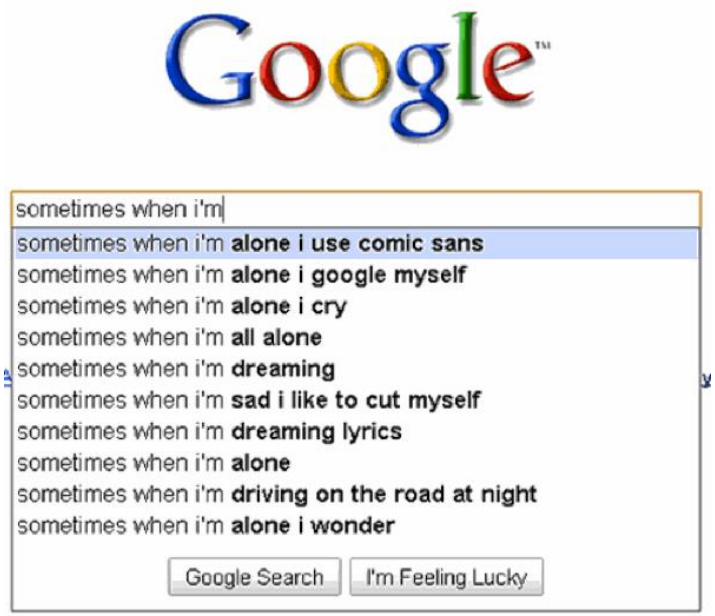
# What's wrong with the current Web?

# The current Web is fantastic!

# Google is also pretty great



# Problem with the current Web: Literature veterans

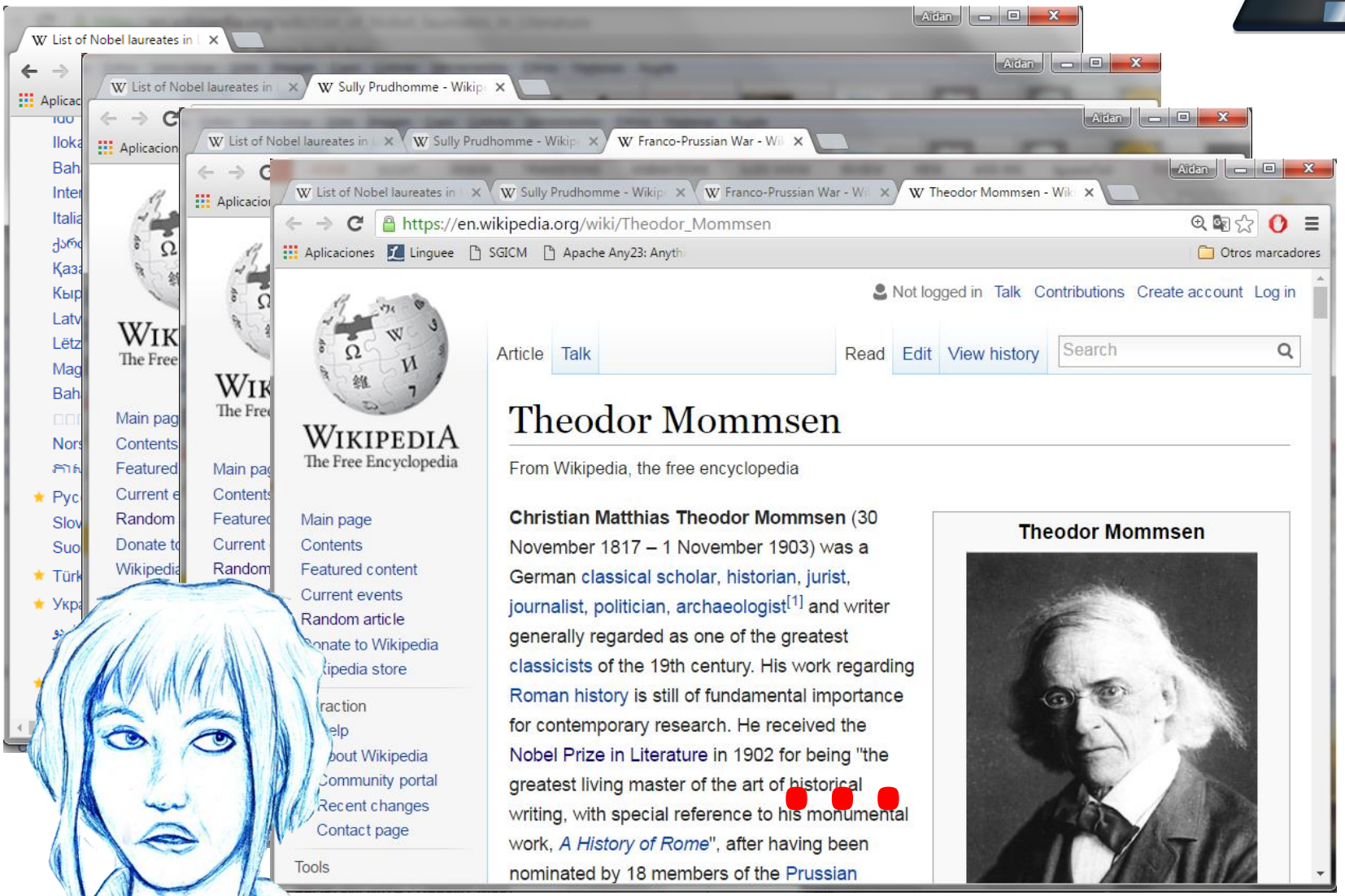
# Doing a report for university …


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 ...



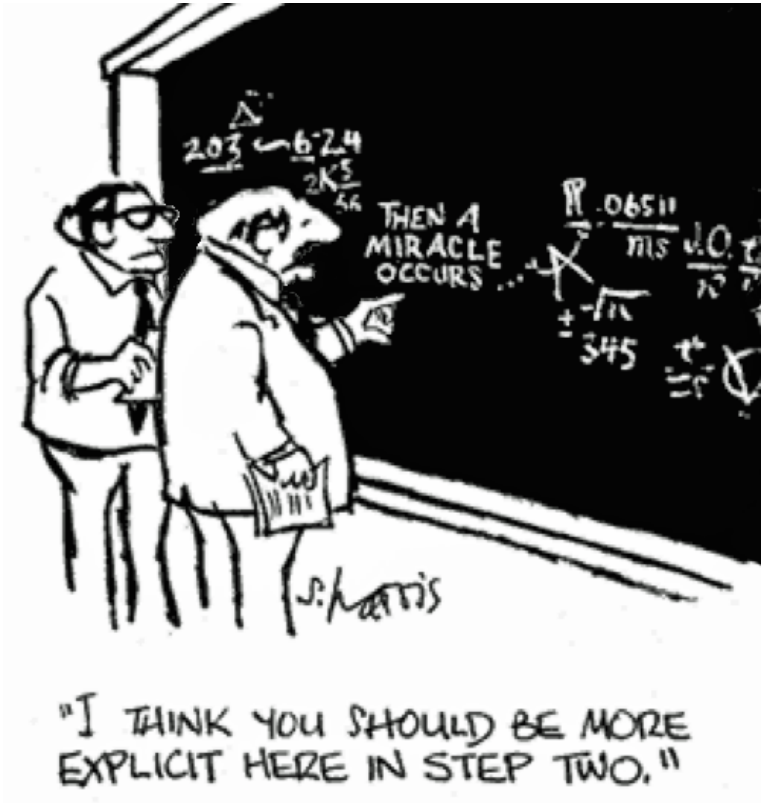
# The Laureate Algorithm (3 decades on)

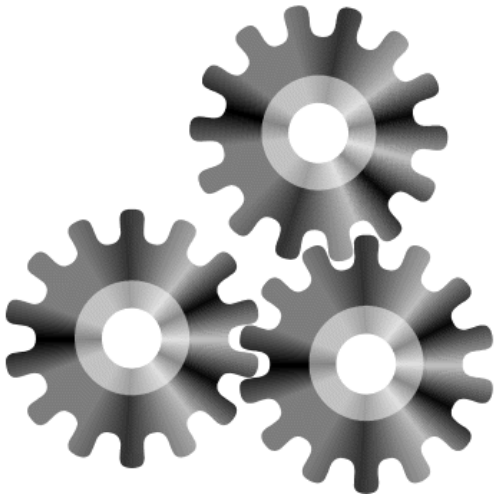
```
results := ∅
for all nobel-lit-winner in wiki-list
    year-prize := year of nobel-lit-winner
    wars := search "war", "conflict", "battle" in nobel-lit-winner
    for all war in wars
            year-war := starting year of war
            add nobel-lit-winner, year-prize, war, year-war to results
    end
end
return results
```

# The Laureate Algorithm (6 decades on?)

*nobel-lit-winner-and-wars* := *magical-sem-web-results()*

return *nobel-lit-winner-and-wars*



# Doing a report for university …


Wants to find all:

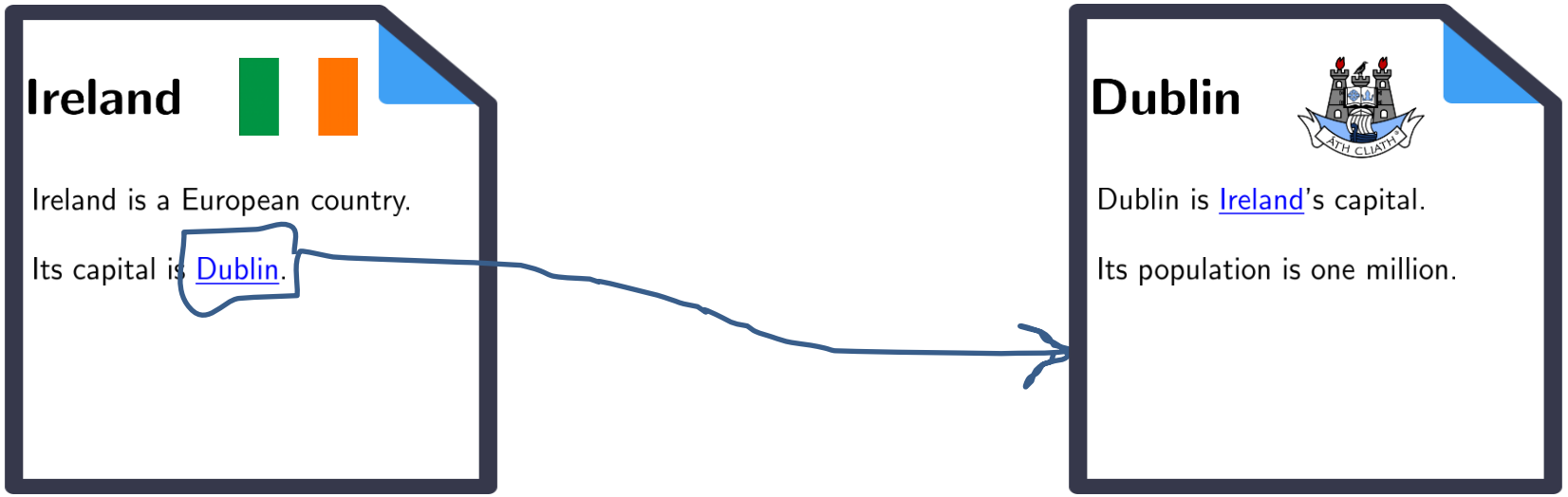
- Nobel Prize winners in Literature
- Who fought in at least one war
- The year they won the prize
- And the year the war(s) started

… why is this query hard on the current Web?

# So what's the problem ...

# The current Web is document-centric

`http://ex.org/Ireland`




`http://ex.org/Dublin`

```
<html>


</html>
```

# The current Web is document-centric

http://ex.org/Ireland


http://ex.org/Dublin


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

# (Most of it) Makes sense to humans

`http://ex.org/Ireland`

**Ireland**

Ireland is a European country.

Its capital is Dublin.

`http://ex.org/Dublin`

**Dublin**

Dublin is Ireland's capital.

Its population is one million.

# ... ASSUMING THEY SPEAK THE LANGUAGE

http://ex.org/Ireland

**Ireland**

Ireland is a European country.

Its capital is Dublin.

http://ex.cl/Dublín

**Dublín**

Dublín es la capital de Irlanda.

Su población es un millón.

# Even worse for machines

`http://ex.org/Ireland`

**Ireland**

Ireland is a European country.

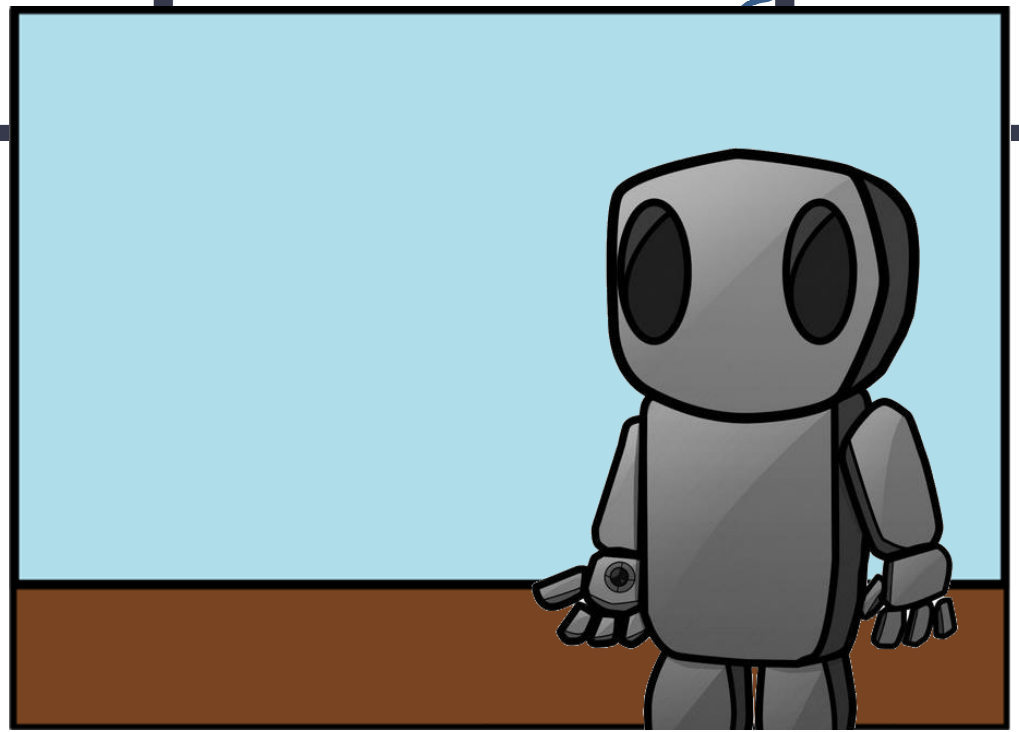
Its capital is Dublin.

`http://ex.cl/Dublín`

**Dublín**

Dublín es la capital de Irlanda.

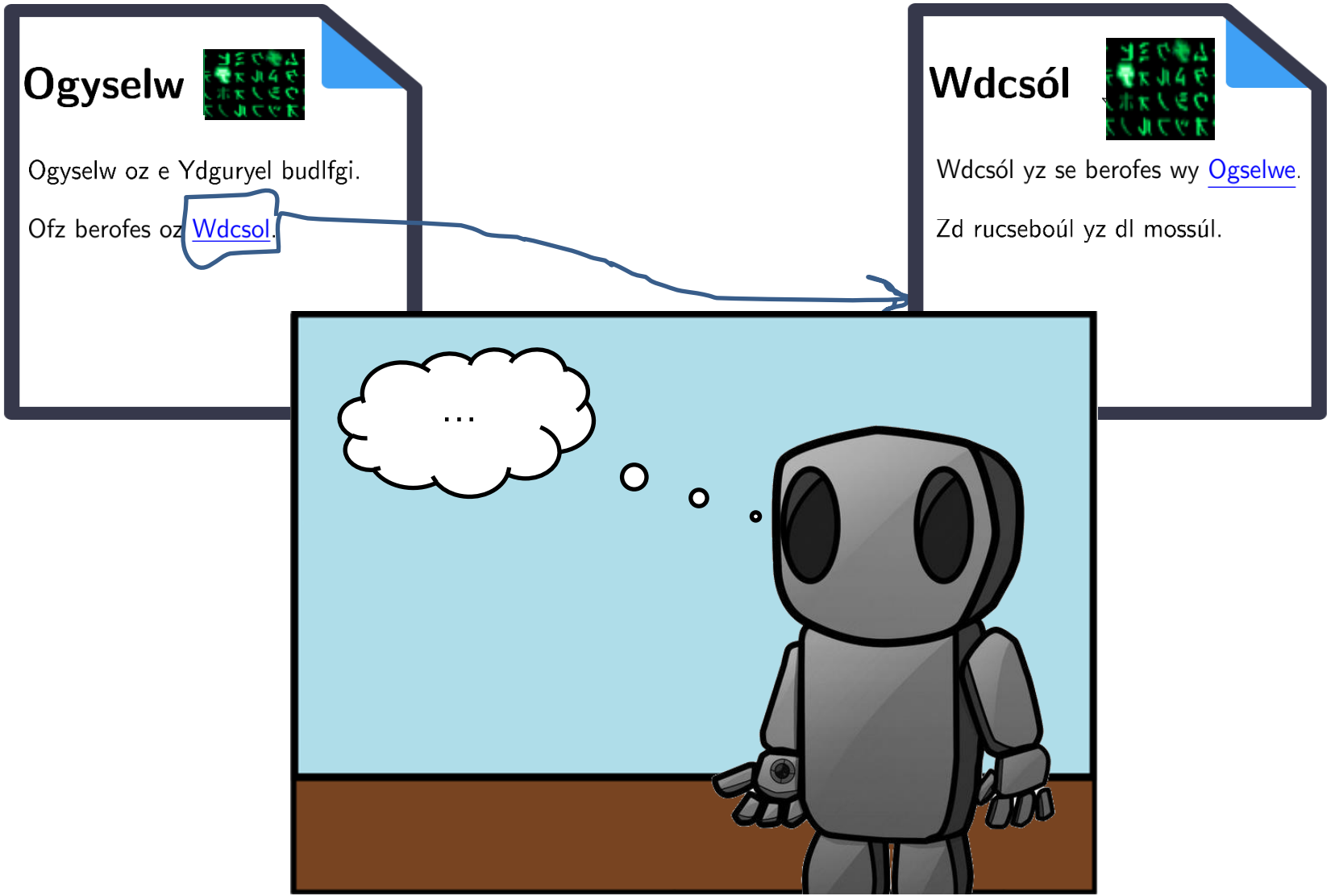
Su población es un millón.



# Ytyl nugzy pug mebholyz

http://ex.org/Ogyselw

http://ex.cl/Wdcsól



# So what's the problem ...

# What machines can do: Fetch documents

http://ex.org/

http://ex.org/Ireland

http://

**Ireland**

Ireland is a European country.

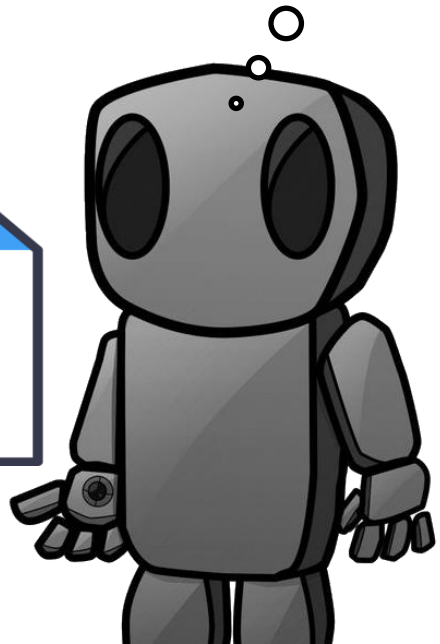
Its capital is Dublin.

HTML

# What machines can do: Find documents





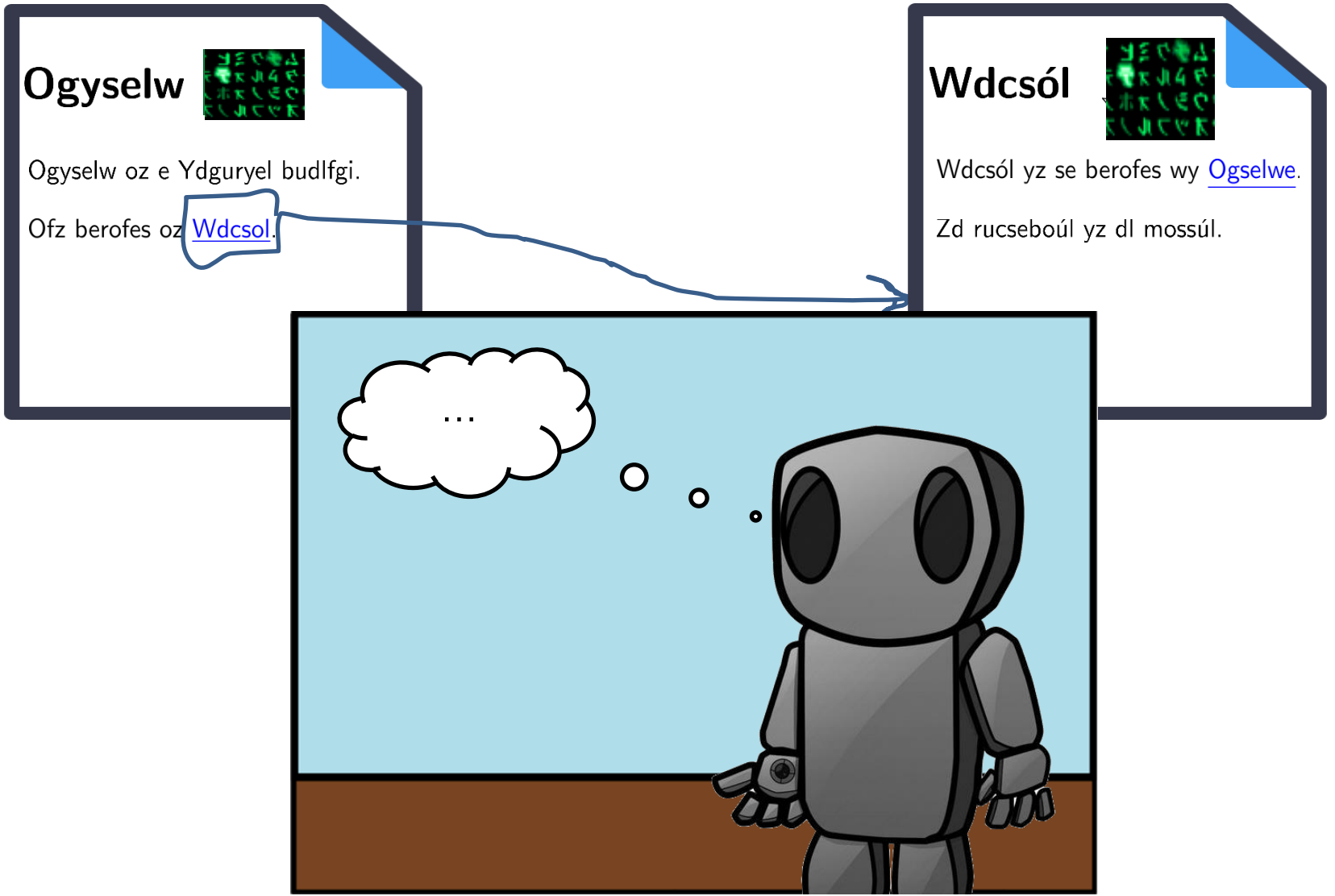


ireland ogyselw

# What machines cannot do: Combine sources

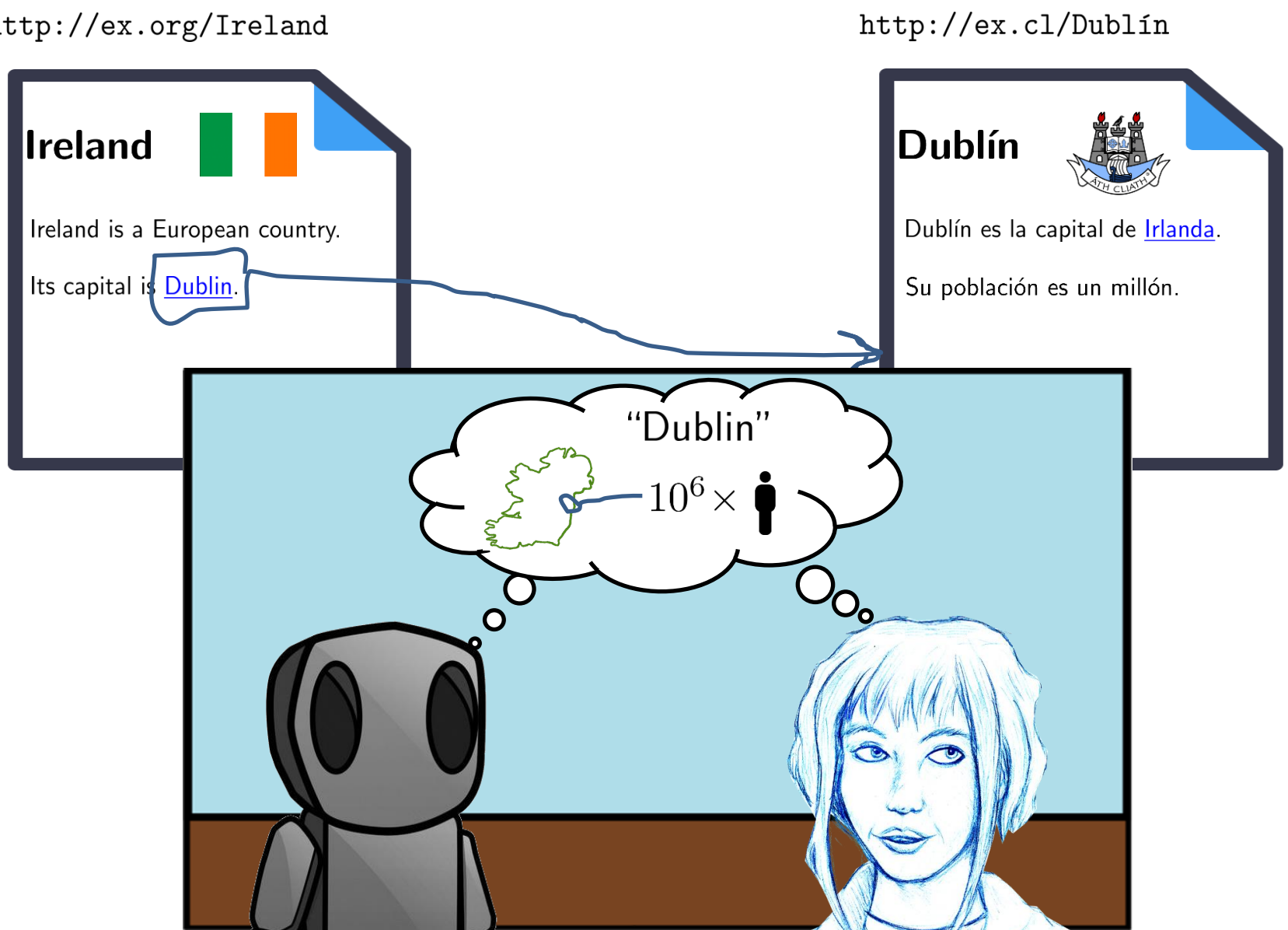
http://ex.org/Ogyselw

http://ex.cl/Wdcsól

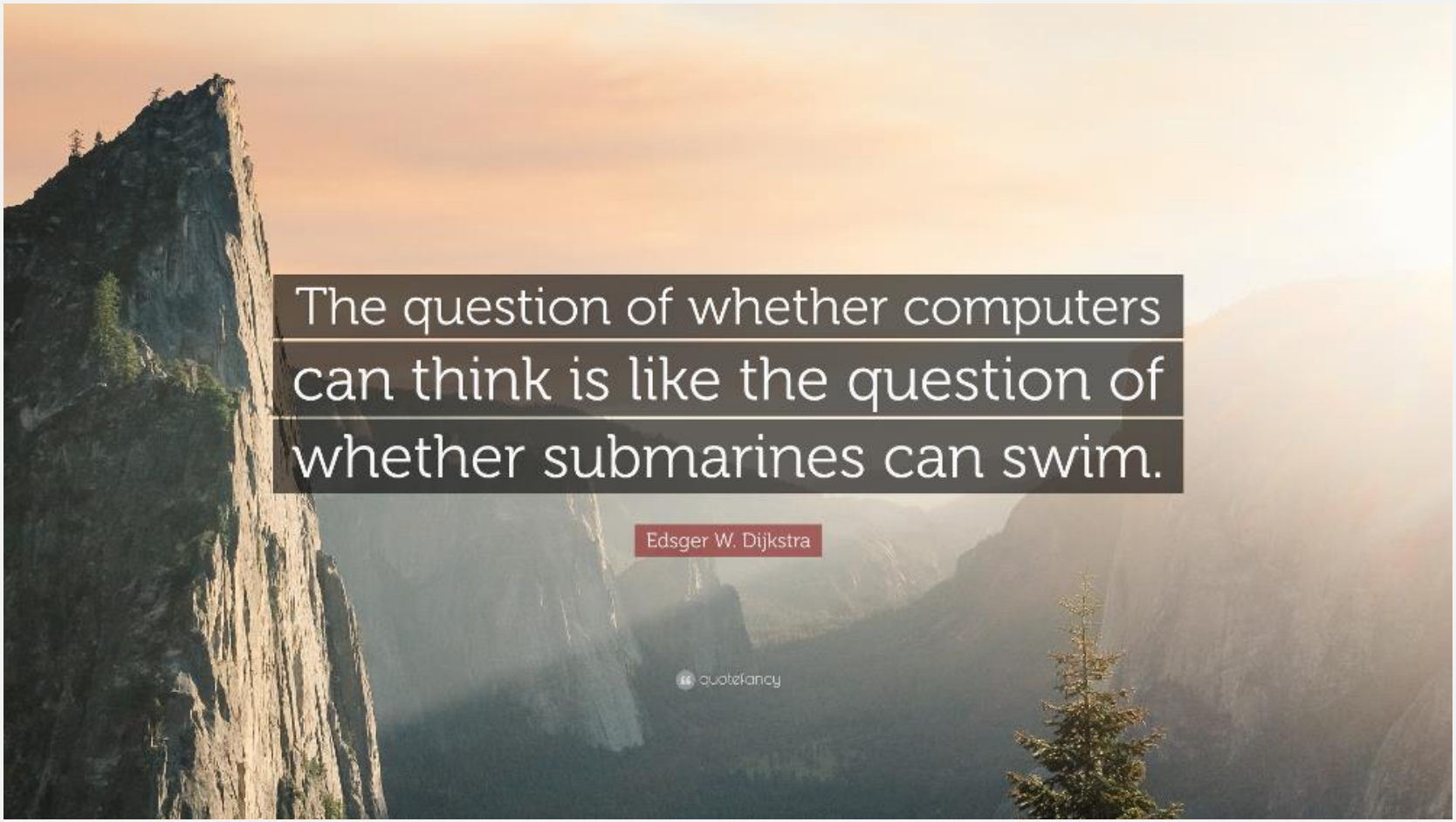


# Towards a semantic web

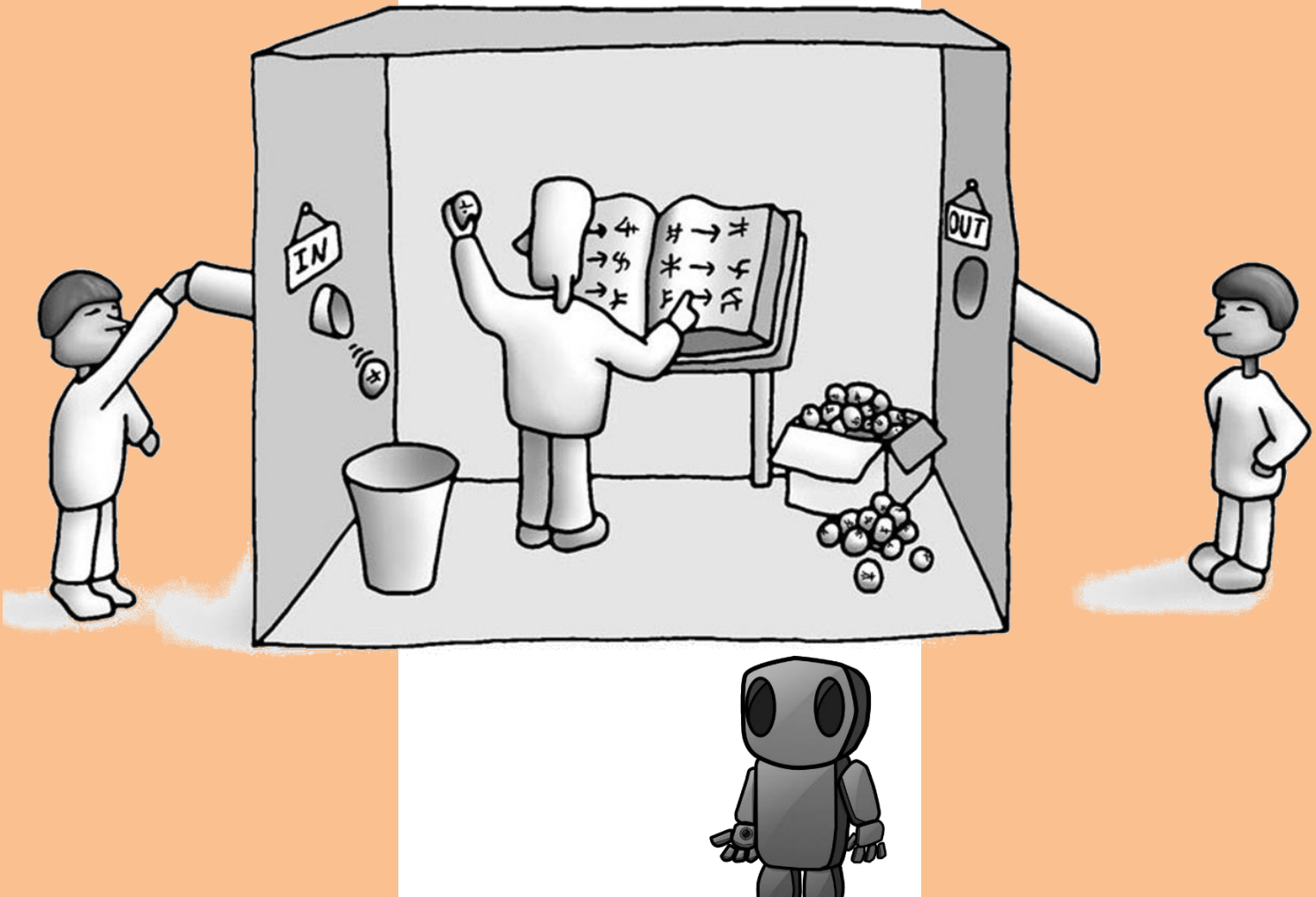
# Machines that "understand" the Web?



# Edsger W. Dijkstra



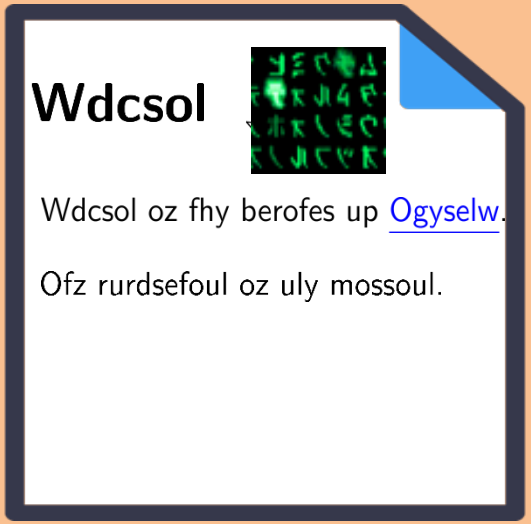
# Searle's Chinese Room ...



# Searle's Chinese Room (Natural Language)

`http://ex.org/Ogyselw`

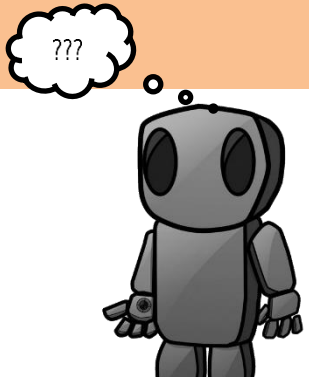

`http://ex.org/Wdcsol`



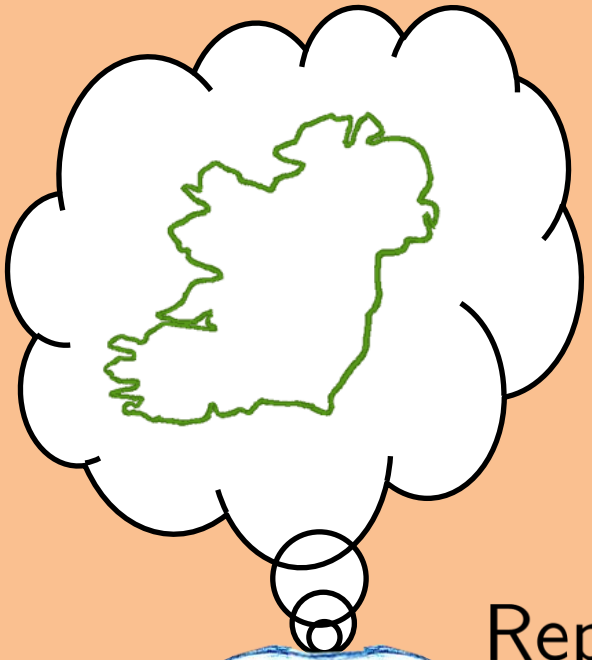

Input: "Nhef oz fhy rurdsefoul up fhy berofes up Ogyselw?"

... what should the output be?

Output: "uly mossoul"



# Multiple names, one thing ...

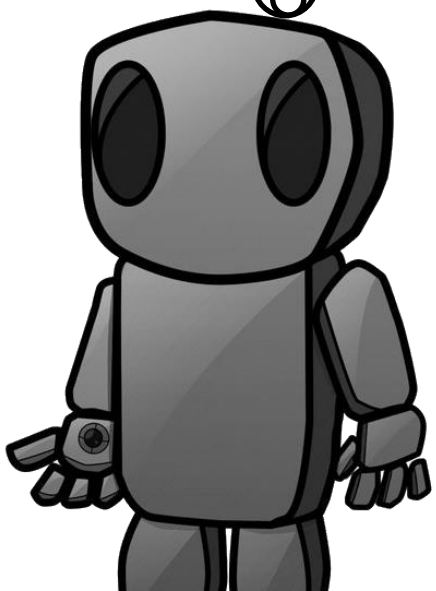


Ireland

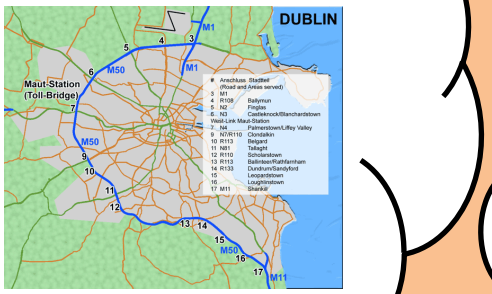
IE

Irlanda

Rep. of Ireland

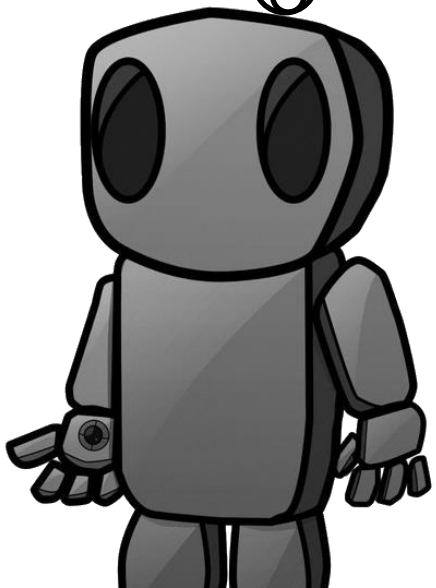


# One name, multiple things …





Dublín



# 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 ...

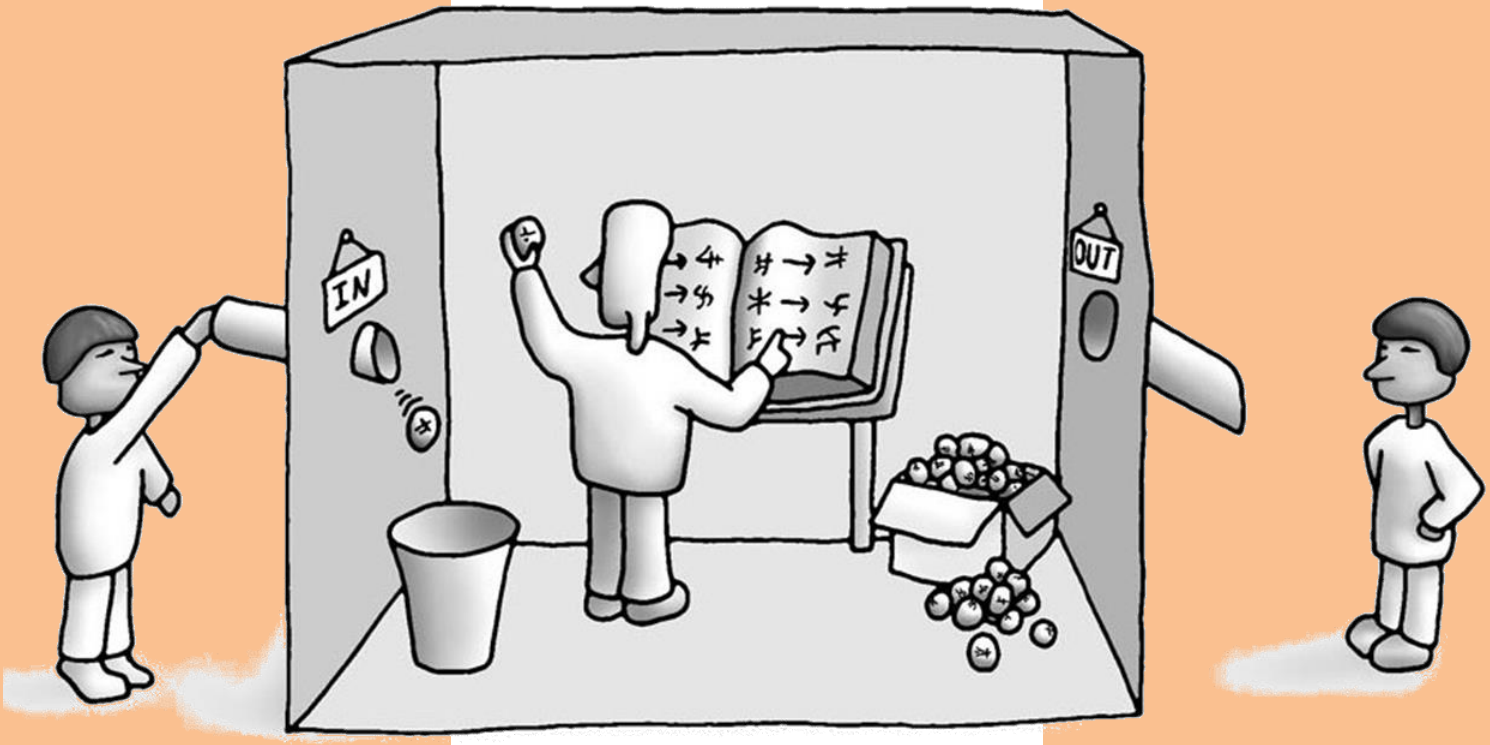
Sherlock saw the man using binoculars.

# Not saying what is meant ...

Fred está arriba la pelota.


# The Semantic Gap

# What if we could "structure" everything ...



# One symbol, one meaning ...

Ireland

IE

Irlanda

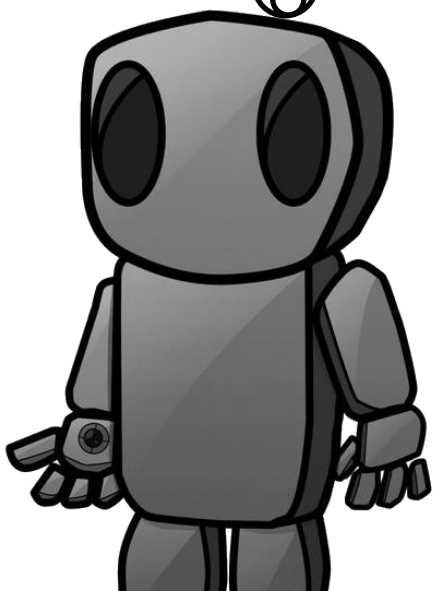
Rep. of Ireland

♇

☿

☾

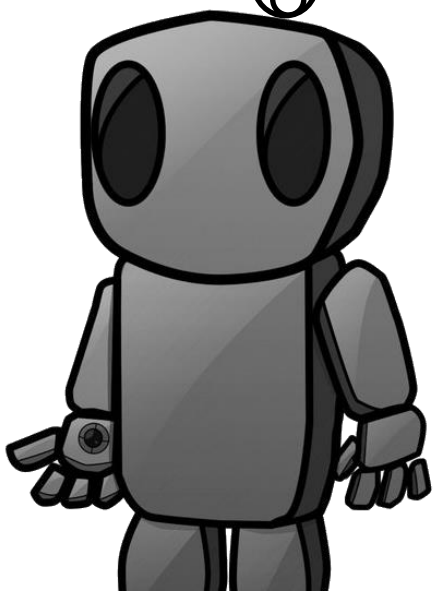
♄



# One symbol, *one* meaning …




Dublín



# 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.

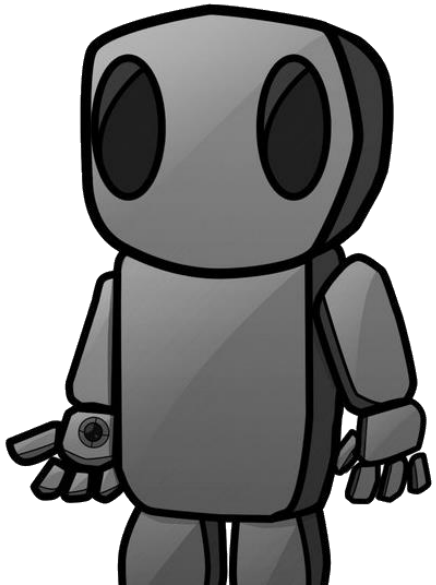
One million people live in Dublin.

[Dublin] Its population is one million.

La población de Dublín es un millón.

(Dublin,population,1000000)

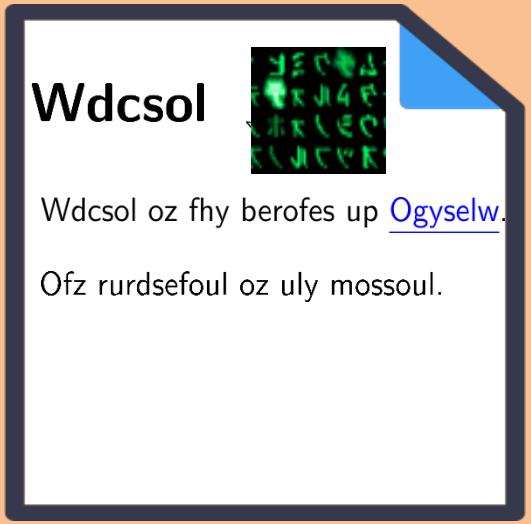
(♅,♂,1000000)



# Searle's Chinese Room (Natural Language)

`http://ex.org/Ogyselw`

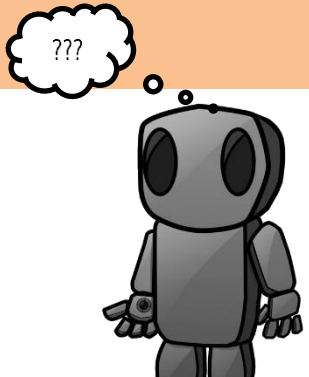

`http://ex.org/Wdcsol`




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

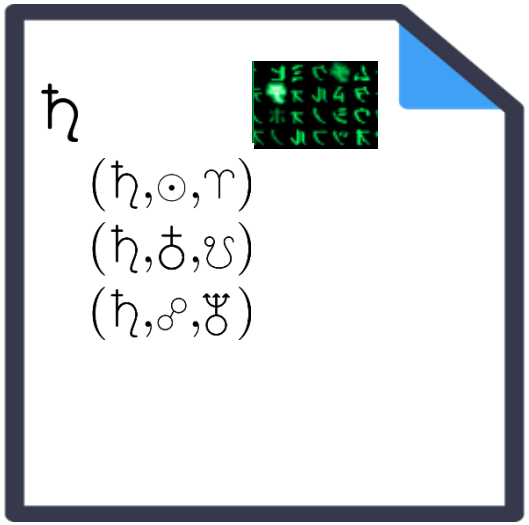
... what should the output be?

Output: "uly mossoul"

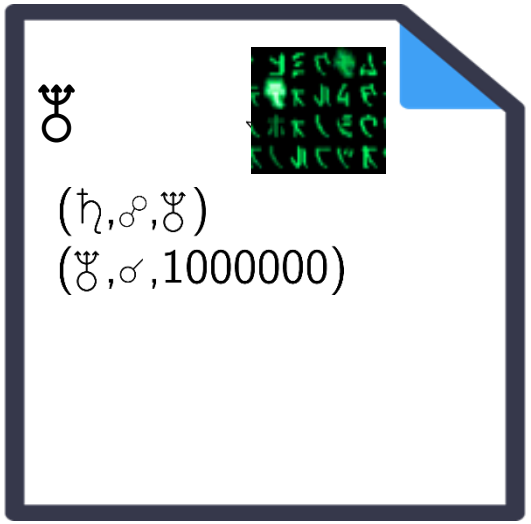


# Searle's Chinese Room (Symbolic)

http://ex.org/♄




http://ex.org/♅




Input: "$(\text{♄},\text{♂},x)$, $(x,\text{♂},y)$?"

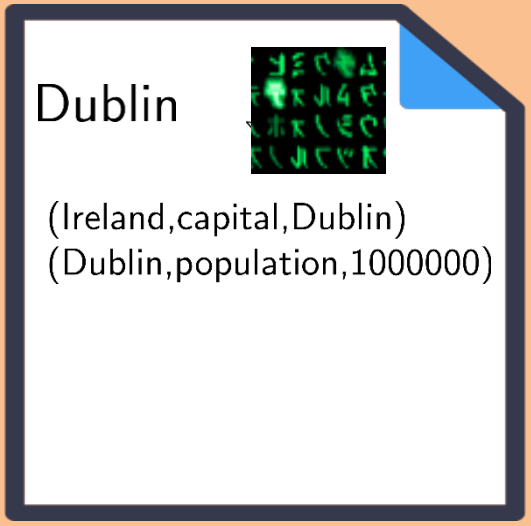
... what should the output be?

Output: $\{(x \mapsto \text{♅}, y \mapsto 1000000)\}$

# Searle's Chinese Room (Symbolic)

`http://ex.org/Ireland`


`http://ex.org/Dublin`



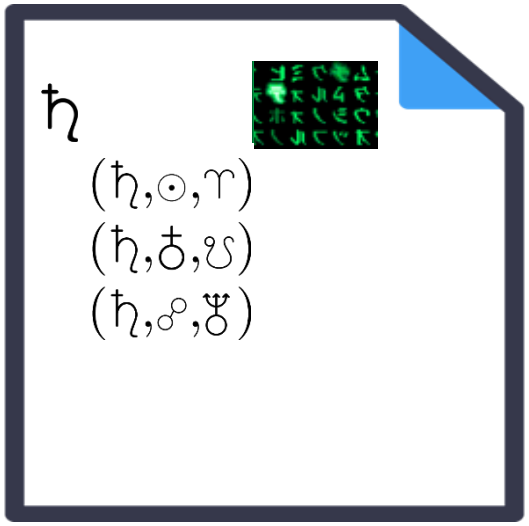

Input: "(Ireland,capital,$x$), ($x$,population,$y$)?"

... what should the output be?

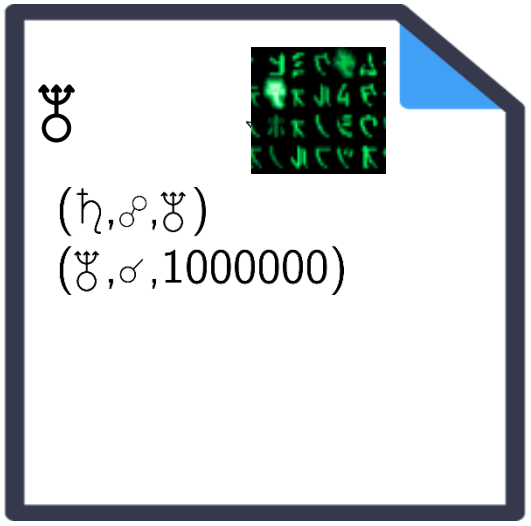
Output: $\{(x \mapsto \text{Dublin}, y \mapsto 1000000)\}$

# Searle's Chinese Room II

http://ex.org/♄




http://ex.org/♆




Input: "$(x,\odot,y)$?"

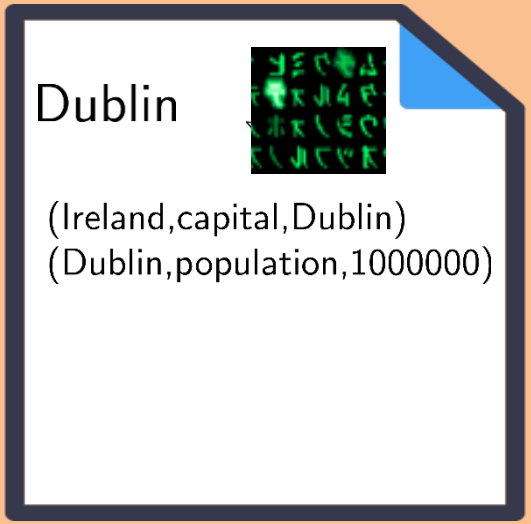
... what should the output be?

Output: $\{(x \mapsto \text{♄}, y \mapsto \text{♈})\}$

# Searle's Chinese Room II

`http://ex.org/Ireland`


`http://ex.org/Dublin`




Input: "($x$,partOf,$y$)?"

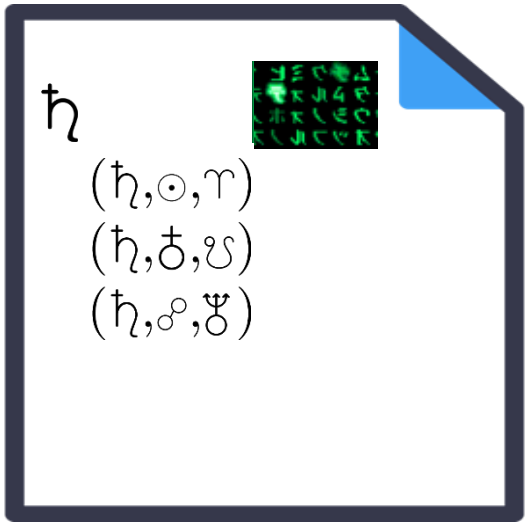
... what should the output be?

Output: $\{(x \mapsto \text{Ireland}, y \mapsto \text{Europe})$

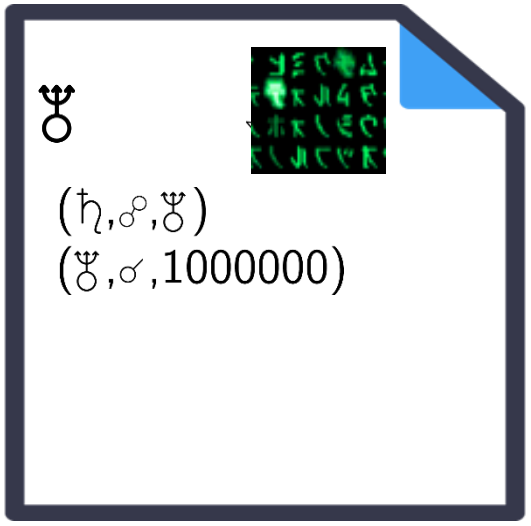
$(x \mapsto \text{Dublin}, y \mapsto \text{Ireland})?$

# Searle's Chinese Room II

http://ex.org/♄




http://ex.org/♆




Rule: "$(b,\male,a) \to (a,\odot,b)$"

Input: "$(x,\odot,y)$?"

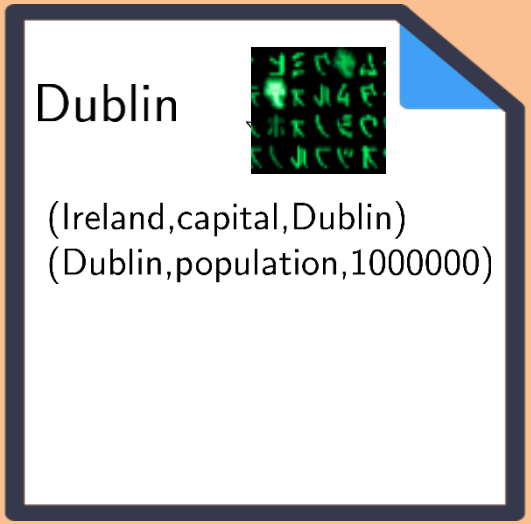
... what should the output be?

Output: $\{(x \mapsto ♄, y \mapsto ♈), (x \mapsto ♆, y \mapsto ♄)\}$

# Searle's Chinese Room II

`http://ex.org/Ireland`


`http://ex.org/Dublin`




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

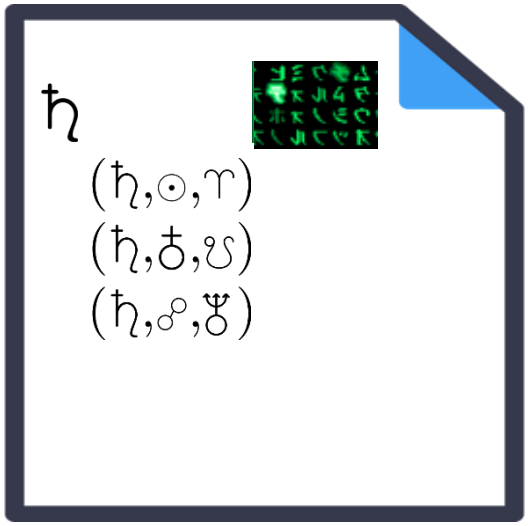
Input: "$(x$,partOf,$y)$?"

... what should the output be?

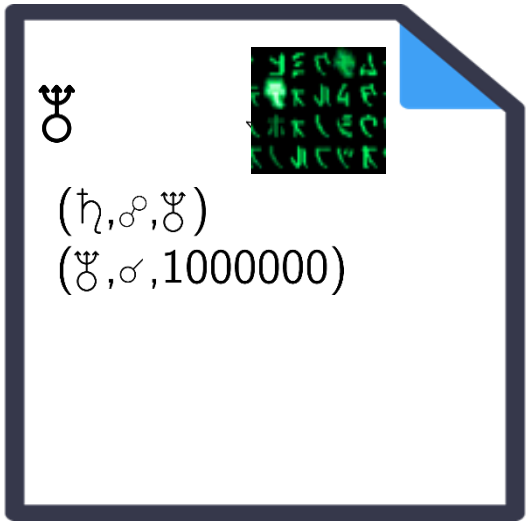
Output: $\{(x \mapsto$ Ireland, $y \mapsto$ Europe),
$(x \mapsto$ Dublin, $y \mapsto$ Ireland$)\}$
$(x \mapsto$ Dublin, $y \mapsto$ Europe$)$?

# Searle's Chinese Room II

http://ex.org/♄




http://ex.org/♅




Rules: "$(b,\male,a) \rightarrow (a,\odot,b)$"

"$(a,\odot,b),\ (b,\odot,c) \rightarrow (a,\odot,c)$"

Input: "$(x,\odot,y)$?"

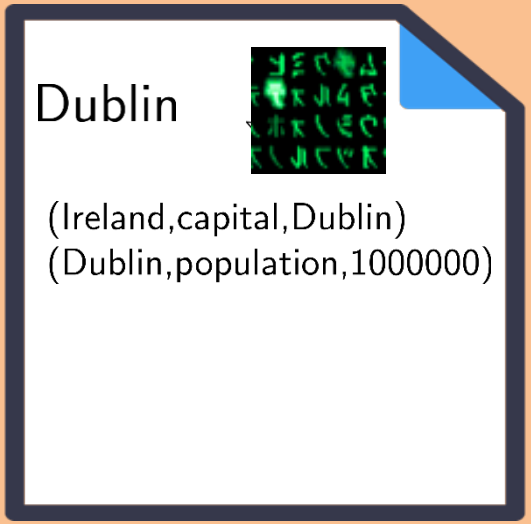
... what should the output be?

Output: $\{(x \mapsto ♄, y \mapsto ♈), (x \mapsto ♅, y \mapsto ♄), (x \mapsto ♅, y \mapsto ♈)\}$

# Searle's Chinese Room II

`http://ex.org/Ireland`


`http://ex.org/Dublin`




Rules: "$(b$,capital,$a)$ → $(a$,partOf,$b)$"
"$(a$,partOf,$b)$, $(b$,partOf,$c)$ → $(a$,partOf,$c)$"

Input: "$(x$,partOf,$y)$?"

... what should the output be?

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

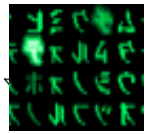
* More or less

# Semantic Web: Data, Logic, Query

Data:

Ireland 

(Ireland,partOf,Europe)
(Ireland,isA,Country)
(Ireland,capital,Dublin)

Dublin 

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

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

Query: "$(x$,partOf,$y)$?"

Output: $\{(x \mapsto$ Ireland, $y \mapsto$ Europe),
$(x \mapsto$ Dublin, $y \mapsto$ Ireland),
$(x \mapsto$ Dublin, $y \mapsto$ Europe$)\}$

# The Semantic Web now?

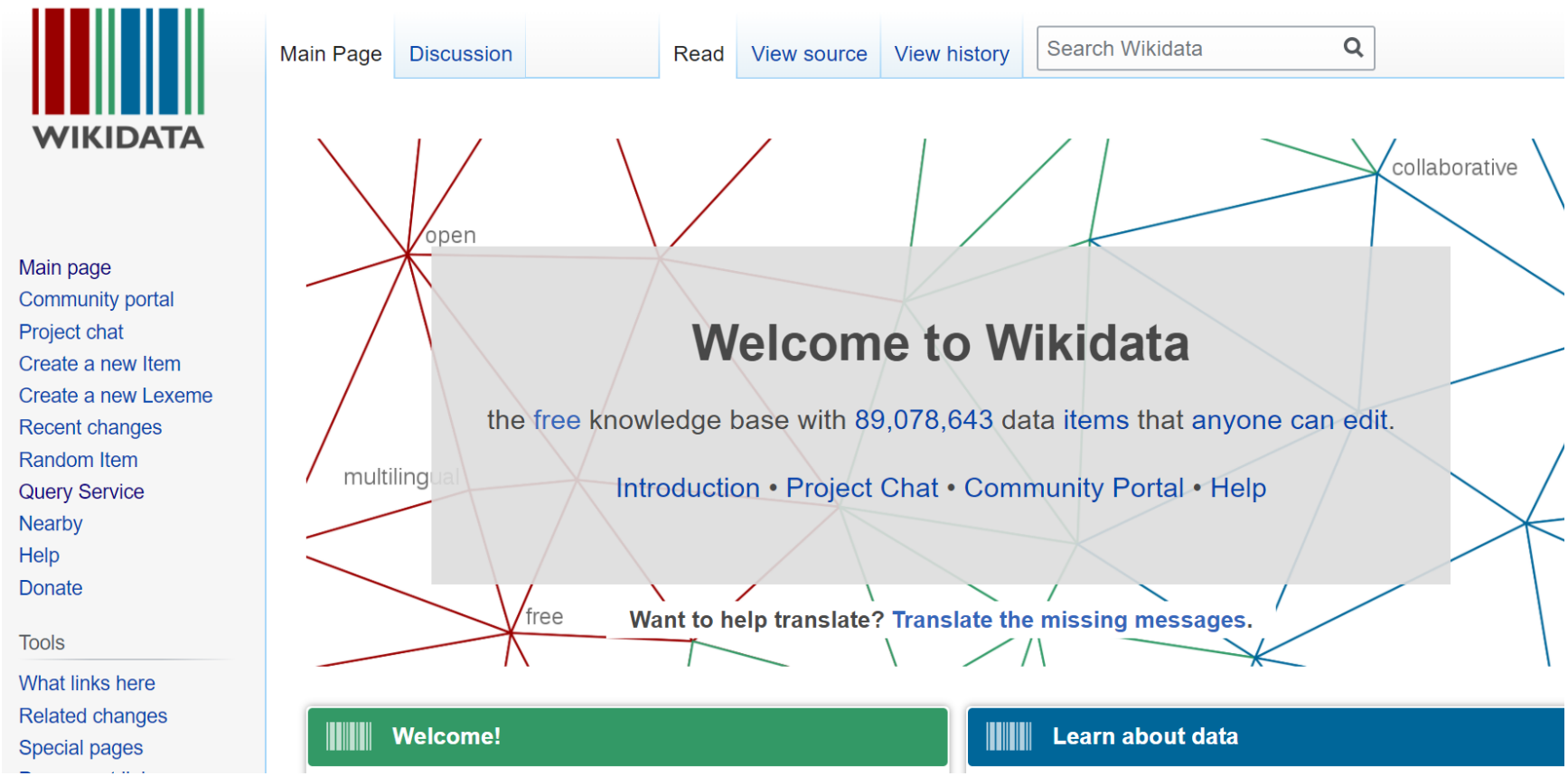
# The Semantic Web is now about 20 years old


… so where is it then?

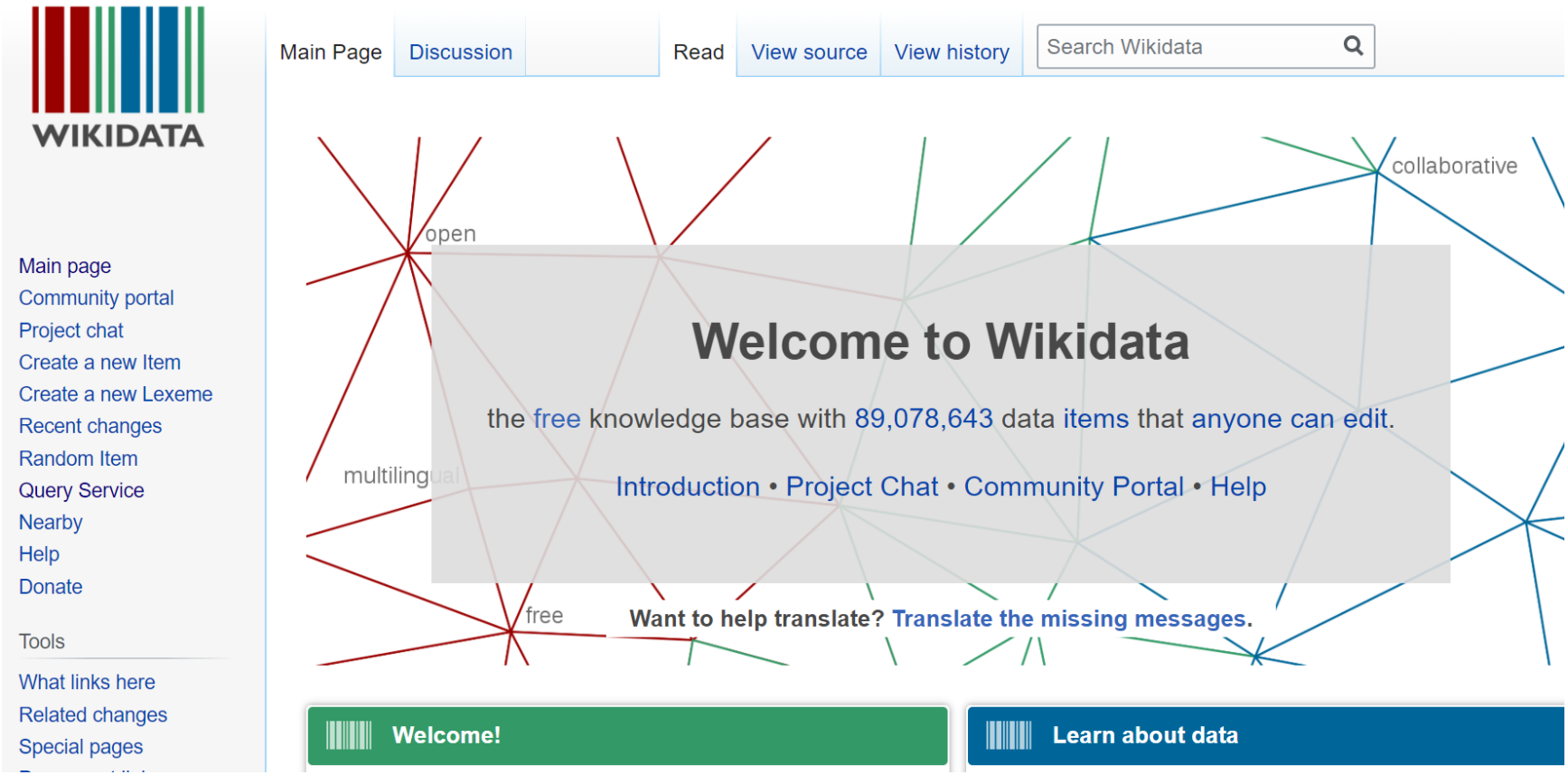
# Hidden within the Web ...

# Wikidata: A Wikipedia for data

# What is Wikidata?



# Why is Wikidata?



# 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][2] |
| **Place of birth** | Tocopilla, Chile[3][2] |

| National team‡ | | | |
|---|---|---|---|
| 2007 | Chile U20 | 12 | (2) |
| 2006– | Chile | 82 | (26) |

**Alexis Sánchez**

| Datos personales | |
|---|---|
| **Nombre completo** | Alexis Alejandro Sánchez Sánchez |

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

Alexis

Alexis Sánchez im Dezember 2011

| Spielerinformationen | |
|---|---|
| **Voller Name** | Alexis Alejandro Sánchez Sánchez |
| **Geburtstag** | 19. Dezember 1988 |
| **Geburtsort** | Tocopilla, Chile |

| Nationalmannschaft[2] | | |
|---|---|---|
| 2007 | Chile U-20 | |
| 2006– | Chile | 76 (25) |

# Problem 2: Complex lists of things manually edited by users


Not logged in Talk Contributions Create account Log in

Article Talk Read Edit View history Search

## Chile national football team

From Wikipedia, the free encyclopedia

### Most capped players [ edit ]

*As of September 1, 2016*

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

| # | Name | International Career | Caps | Goals |
|---|---|---|---|---|
| 1. | **Claudio Bravo** | 2004 – | 106 | 0 |
| 2. | **Alexis Sánchez** | 2006 – | 102 | 34 |
| 3. | **Gary Medel** | 2007 – | 96 | 7 |
| 4. | **Gonzalo Jara** | 2006 – | 95 | 3 |

### Top goalscorers [ edit ]

*As of September 1, 2016*

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

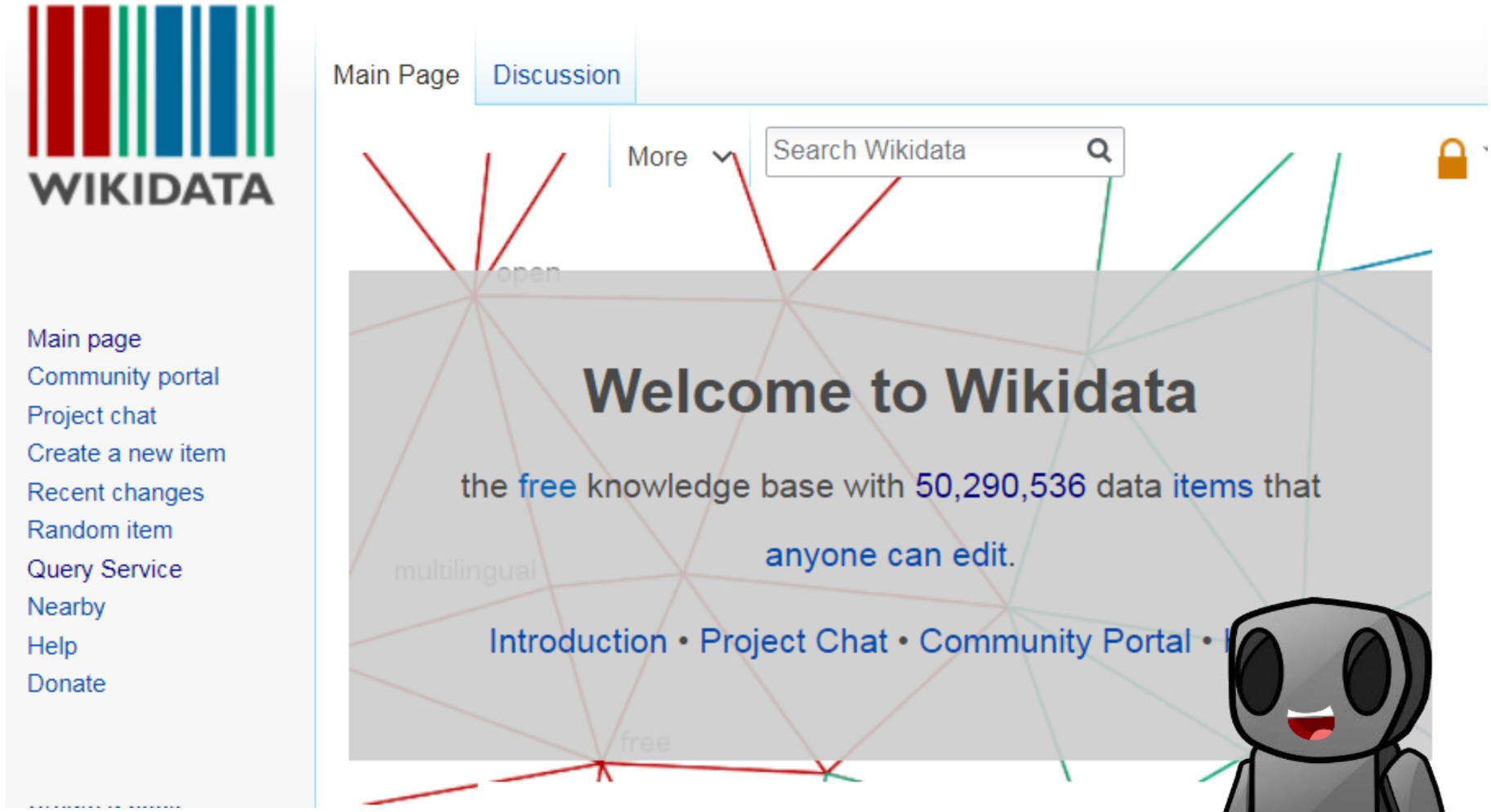
| # | Name | International Career | Goals | Caps |
|---|---|---|---|---|
| 1. | Marcelo Salas | 1994–2007 | 37 | 70 |
| 2. | Iván Zamorano | 1987–2001 | 34 | 69 |
| 2. | **Alexis Sánchez** (list) | 2006 – | 34 | 102 |

# Alexis scores a goal …


Now an army of human editors has to manually update a bunch of articles:

different languages, lists, …

# Solution: Wikidata





# Alexis Alejandro Sánchez Sánchez (Q180553)

Chilean footballer

Chiqui

edit

In more languages

## Statements

| | | |
|---|---|---|
| sex or gender | male | edit |
| | 1 reference | |
| | + add | |
| member of sports team | Cobreloa | edit |
| | 1 reference | |
| | Udinese Calcio | edit |
| | 1 reference | |
| | Arsenal F.C | edit |
| | 1 reference | |
| | Club Atlético River Plate | edit |

Main page
Community portal
Project chat
Create a new item
Item by title
Recent changes
Random item
Nearby
Help
Donate

Print/export
Create a book
Download as PDF
Printable version

Tools
What links here
Related changes
Special pages
Permanent link
Page information
Concept URI
Cite this page

# Alexis Sánchez (Q180553)

futbolista chileno editar

Alexis Sánchez Sánchez | AS9 | Alexis Alejandro Sánchez Sánchez | AS7 | Alexis Alejandro Sánchez | Alexis Sánches | Alexis Sanchez | Alexis Sanches

▸ En más idiomas

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 imprimir

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

## Declaraciones

| | | |
|---|---|---|
| sexo | masculino | editar |
| | ▸ 1 referencia | |
| | | + añadir |
| miembro del equipo deportivo | Club de Deportes Cobreloa | editar |
| | ▸ 1 referencia | |
| | Udinese Calcio | editar |
| | ▸ 1 referencia | |
| | Arsenal Football Club | editar |
| | ▸ 1 referencia | |

# Alexis Sánchez (Q180553)

chilenischer Fußballspieler

bearbeiten

Alexis Alejandro Sánchez Sánchez | Alexis Sanchez | Alexis Alejandro Sanchez Sanchez

In weiteren Sprachen

Hauptseite
Gemeinschaftsportal
Forum
Ein neues Datenobjekt erstellen
Datenobjekte nach Titel
Letzte Änderungen
Zufälliges Datenobjekt
In der Nähe
Hilfe
Spenden

Drucken/exportieren
Buch erstellen
Als PDF herunterladen
Druckversion

Werkzeuge
Links auf diese Seite
Änderungen an verlinkten Seiten
Spezialseiten
Permanenter Link
Seiteninformationen
Konzept-URI
Seite zitieren

## Aussagen

| | | |
|---|---|---|
| Geschlecht | männlich<br>1 Fundstelle | bearbeiten |
| | + hinzufügen | |
| Mitglied von Sportmannschaft oder -verein | CD Cobreloa<br>1 Fundstelle | bearbeiten |
| | Udinese Calcio<br>1 Fundstelle | bearbeiten |
| | FC Arsenal<br>1 Fundstelle | bearbeiten |
| | CA River Plate | bearbeiten |

# Санчес, Алексис (Q180553)

Описание не заполнено

редактировать

Алексис Санчес

На других языках

## Утверждения

| | | |
|---|---|---|
| пол | мужской | редактировать |
| | 1 источник | |
| | | + добавить |
| член спортивной команды | Кобрелоа | редактировать |
| | 1 источник | |
| | Удинезе | редактировать |
| | 1 источник | |
| | Арсенал | редактировать |
| | 1 источник | |
| | Ривер Плейт | редактировать |

Заглавная страница
Портал сообщества
Форум
Создание нового элемента
Элементы по заголовку
Свежие правки
Случайный элемент
Рядом
Справка
Пожертвования

Печать/экспорт
Создать книгу
Скачать как PDF
Версия для печати

Инструменты
Ссылки сюда
Связанные правки
Спецстраницы
Постоянная ссылка
Сведения о странице
URI концепта
Цитировать страницу

# അലക്സിസ് സാഞ്ചസ് (Q180553)

വിവരണമൊന്നും നിർവചിച്ചിട്ടില്ല
അപരനാമങ്ങളൊന്നും കണ്ടെത്താനായില്ല.

തിരുത്തുക

കൂടുതൽ ഭാഷകളിൽ

പ്രധാന താൾ
സാമൂഹികകവാടം
Project chat
പുതിയൊരു ഇനം സൃഷ്ടിക്കുക
ഇനം തലക്കെട്ടനുസരിച്ച്
സമീപകാല മാറ്റങ്ങൾ
ഏതെങ്കിലും താൾ
സമീപസ്ഥം
സഹായം
സംഭാവന

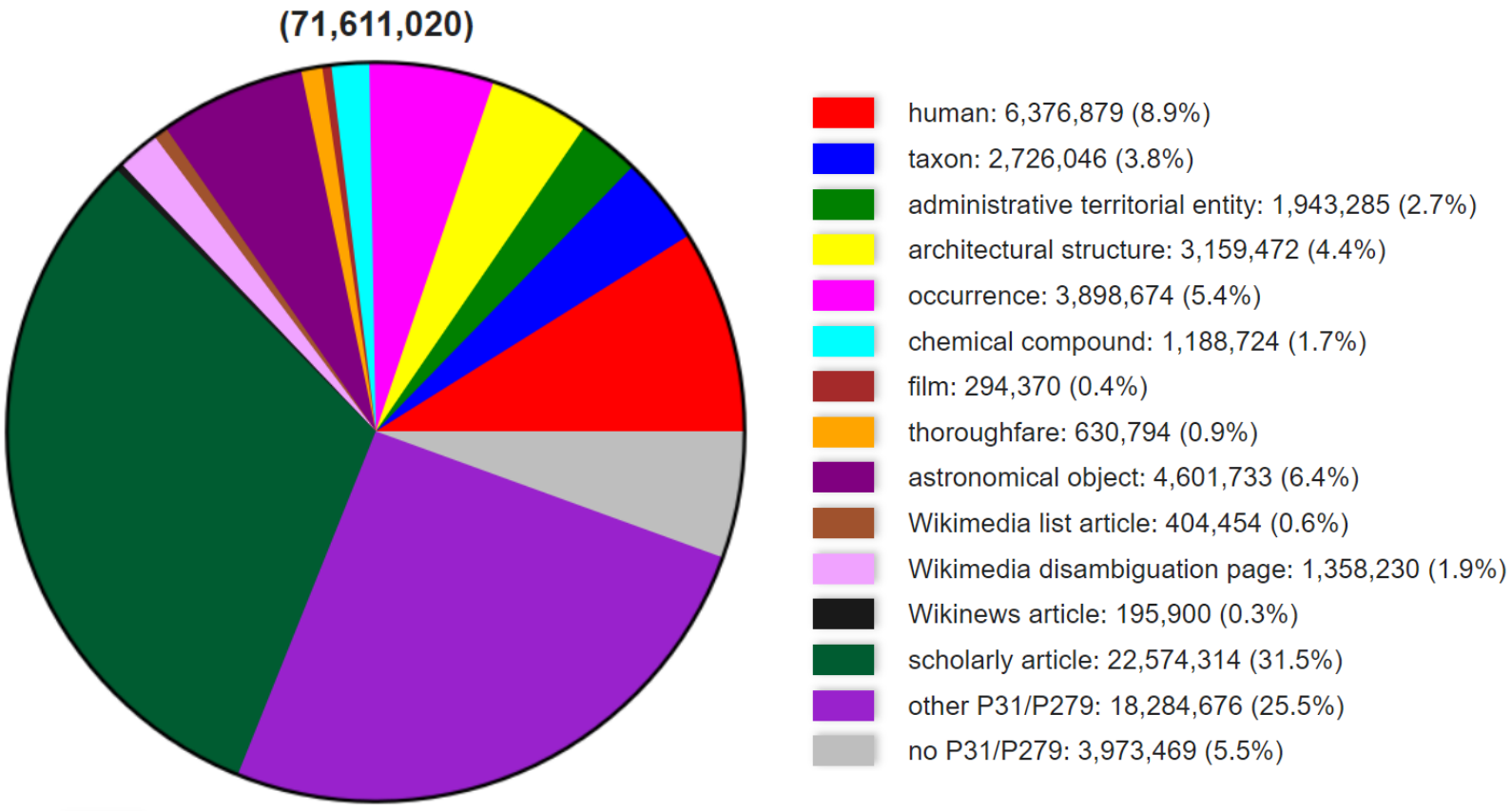
അച്ചടിയ്ക്കുക/കയറ്റുമതി ചെയ്യുക
പുസ്തകം സൃഷ്ടിക്കുക
PDF ആയി ഡൗൺലോഡ് ചെയ്യുക
അച്ചടിരൂപം

ഉപകരണങ്ങൾ
ഈ താളിലേക്കുള്ള കണ്ണികൾ
അനുബന്ധ മാറ്റങ്ങൾ
പ്രത്യേക താളുകൾ
സ്ഥിരംകണ്ണി

## Statements

| | | |
|---|---|---|
| ലിംഗം | പുരുഷൻ<br>1 സ്രോതസ്സ് | തിരുത്തുക |
| | | + ചേർക്കുക |
| member of sports team *ഇംഗ്ലീഷ്* | Cobreloa *ഇംഗ്ലീഷ്*<br>1 സ്രോതസ്സ് | തിരുത്തുക |
| | Udinese Calcio *ഇംഗ്ലീഷ്*<br>1 സ്രോതസ്സ് | തിരുത്തുക |
| | ആഴ്സണൽ എഫ്.സി.<br>1 സ്രോതസ്സ് | തിരുത്തുക |
| | Club Atlético River Plate *ഇംഗ്ലീഷ്* | തിരുത്തുക |

# What does Wikidata describe?




- 23,593 active users

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

# Use-case: Info-Boxes

WIKIPEDIA
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**

The APEX telescope

| | |
|---|---|
| **Observatory** | Llano de Chajnantor Observatory |
| **Location(s)** | Atacama Desert, Chile |
| **Coordinates** | 23°00′21″S 67°45′33″W |
| **Organization** | Edit this at Wikidata<br>Observatory<br>Max Planck Institute for Radio Astronomy<br>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<br>Cosmic microwave background experiment<br>Radio telescope |

# Use-case: Info-Boxes

WIKIPEDIA
The Free Encyclopedia

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

Main page
Contents
Featured content
Current events

WIKIDATA

Item Discussion Read View history Search Wikidata

## Atacama Pathfinder Experiment (Q753076)

radio telescope

In more languages

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

### Statements

| | | |
|---|---|---|
| instance of | radio telescope | edit |
| | 0 references | |
| | + add reference | |
| | Cassegrain reflector | edit |
| | 0 references | |
| | + add reference | |

**Atacama Pathfinder Experiment**

The APEX telescope

| | |
|---|---|
| **Observatory** | Llano de Chajnantor Observatory |
| **Location(s)** | Atacama Desert, Chile |
| **Coordinates** | 23°00′21″S 67°45′33″W |
| **Organization** | Edit this at Wikidata<br>Observatory<br>Max Planck Institute for Radio Astronomy<br>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<br>Cosmic microwave background experiment<br>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)

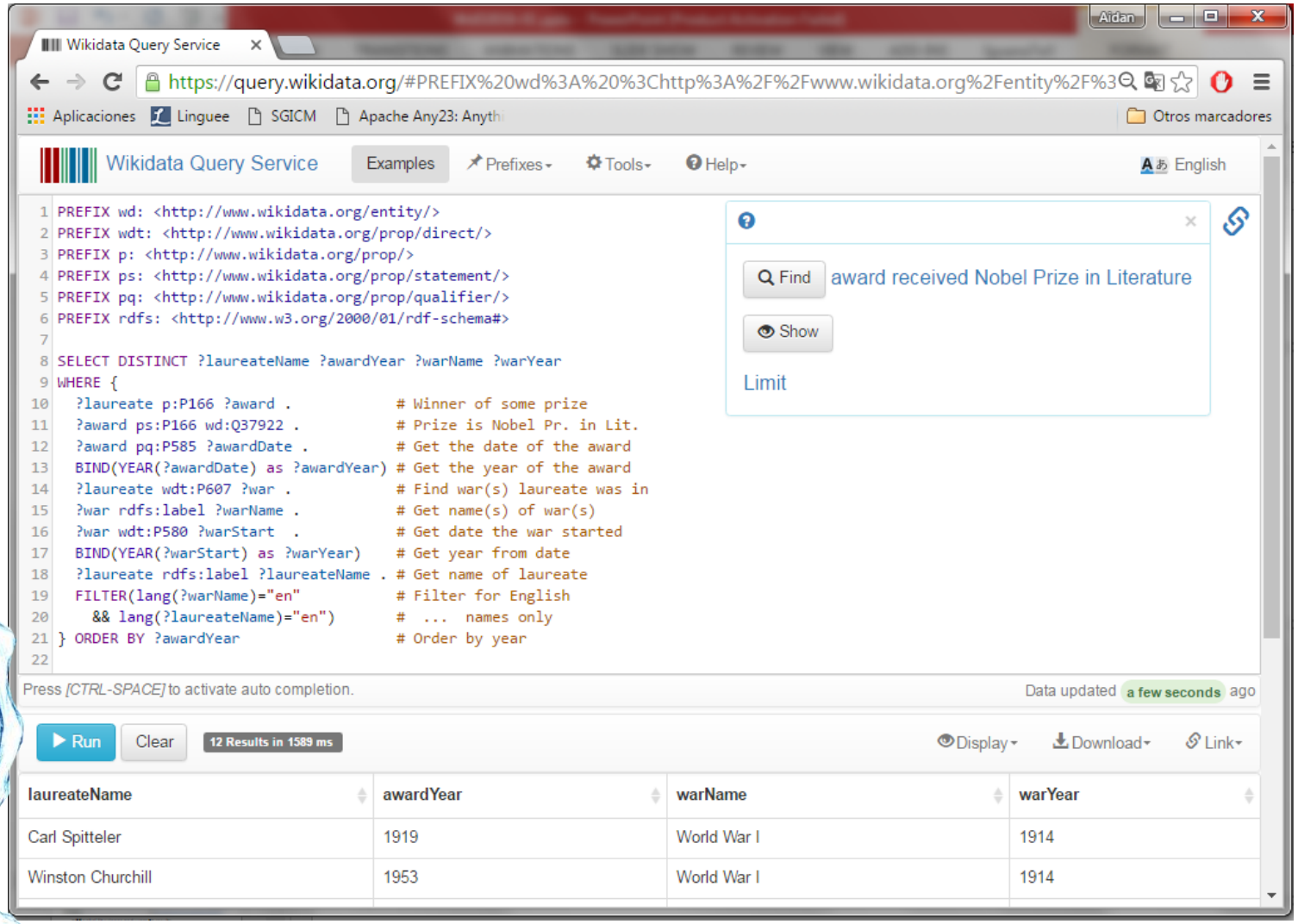
| Item | Birth | Death | Age |
|---|---|---|---|
| Tuti Yusupova (Q1038827) | 1880-07-01 | 2015-03-28 | 134 |
| Karni Mata (Q1106783) | 1387-10-02 | 1538-03-23 | 150 |
| no label (Q11556831) | 1185-01-01 | 1392-01-01 | 207 |
| Minamoto no Chikayuki (Q11564306) | 1185-01-01 | 1333-01-01 | 148 |
| Bir 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 |

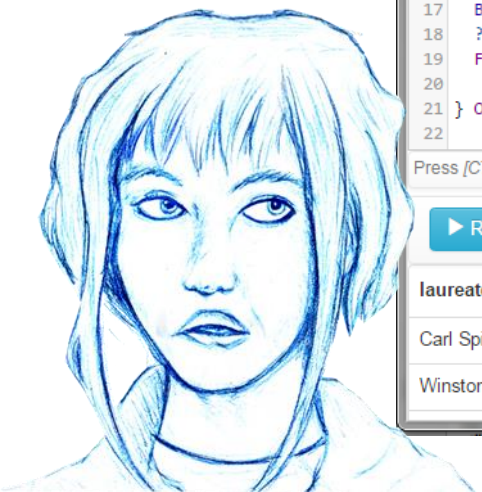
# Use-case: Quality Checks

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

Wikidata Query Service

https://query.wikidata.org/#PREFIX%20wd%3A%20%3Chttp%3A%2F%2Fwww.wikidata.org%2Fentity%2F%3

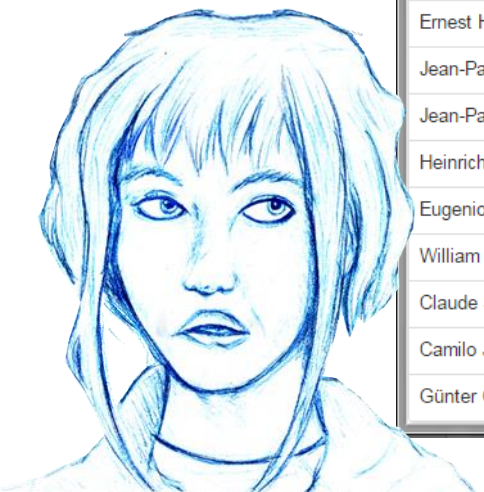
Aplicaciones Linguee SGICM Apache Any23: Anyth Otros marcadores

```
16   ?war wdt:P580 ?warStart  .          # Get date the war started
17   BIND(YEAR(?warStart) as ?warYear)   # Get year from date
18   ?laureate rdfs:label ?laureateName . # Get name of laureate
19   FILTER(lang(?warName)="en"          # Filter for English
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 Display Download Link

| laureateName | awardYear | warName | warYear |
|---|---|---|---|
| Carl Spitteler | 1919 | World War I | 1914 |
| Winston Churchill | 1953 | World War I | 1914 |
| Ernest Hemingway | 1954 | World War I | 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 |



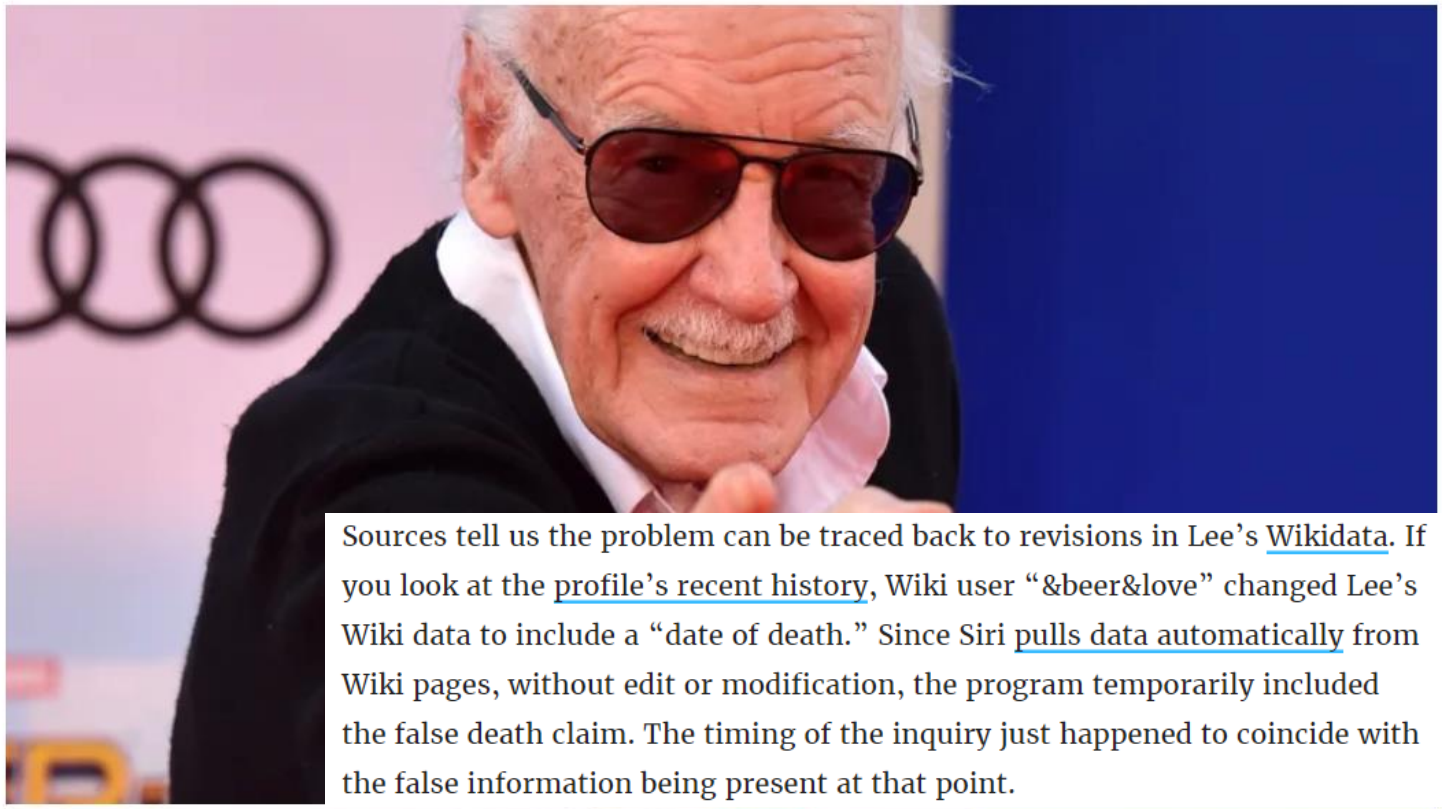
# Used in Applications like Siri ...

## Siri Erroneously Told People Stan Lee Was Dead



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

17.4K 45 5

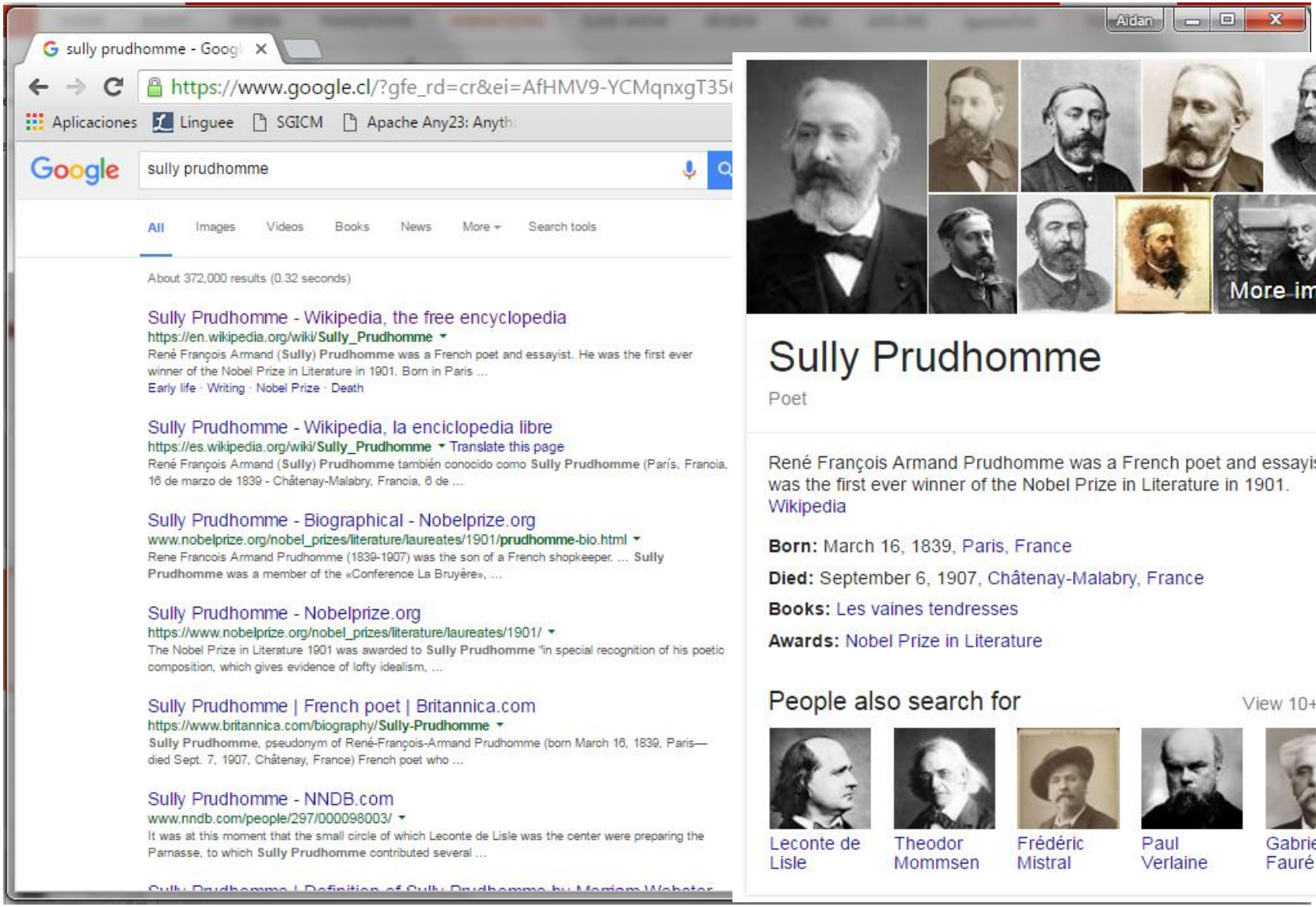


Sources tell us the problem can be traced back to revisions in Lee's Wikidata. If you look at the profile's recent history, Wiki user "&beer&love" changed Lee's Wiki data to include a "date of death." Since Siri pulls data automatically from Wiki pages, without edit or modification, the program temporarily included the false death claim. The timing of the inquiry just happened to coincide with the false information being present at that point.

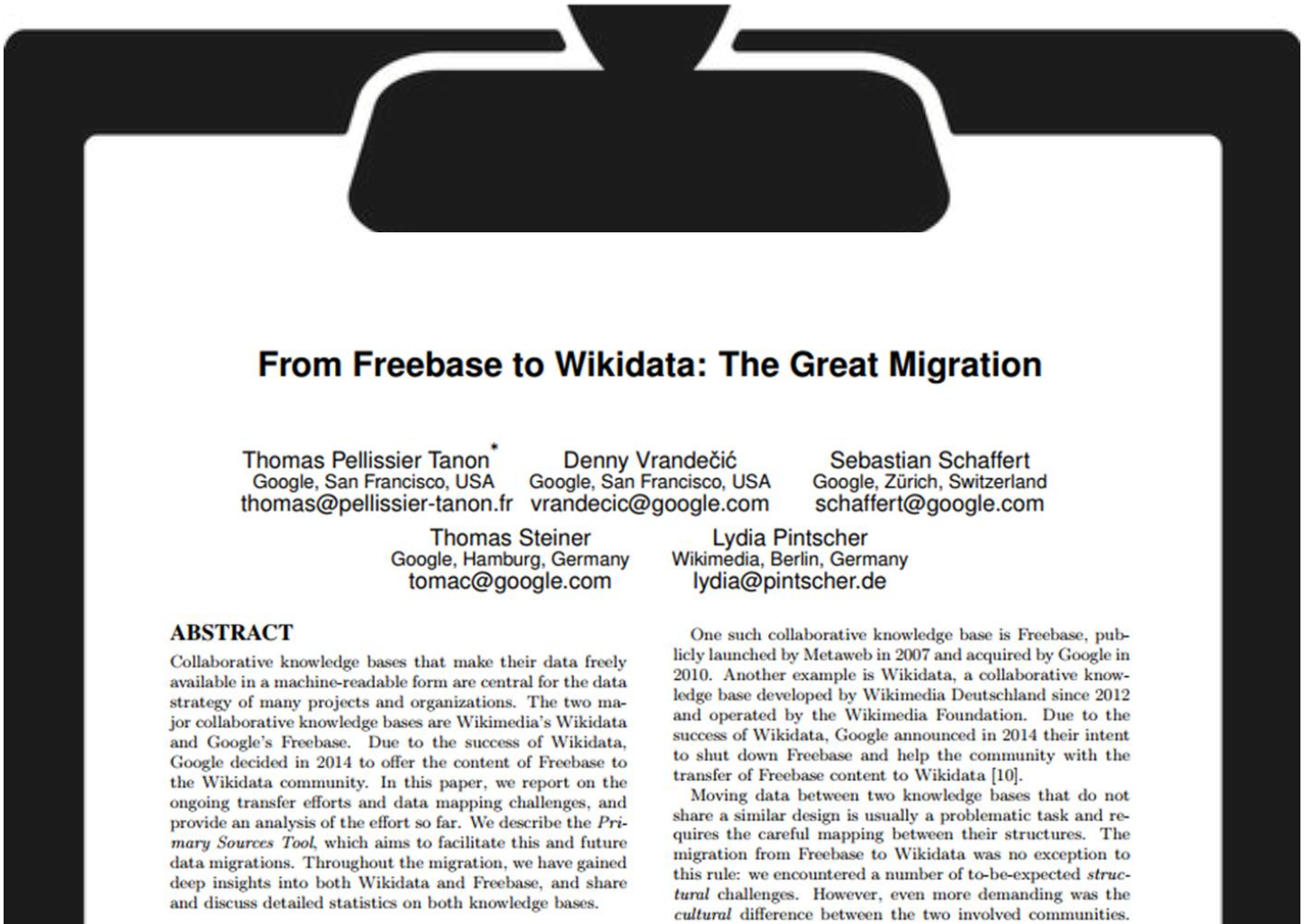
Stan Lee at the premiere of *Spider-Man: Homecoming*.
Photo: Alberto E. Rodriguez (Getty Images)

# Google's Knowledge Graph

# Google's Knowledge Panel



# Using Semantic Web knowledge-bases



## From Freebase to Wikidata: The Great Migration

Thomas Pellissier Tanon*
Google, San Francisco, USA
thomas@pellissier-tanon.fr

Denny Vrandečić
Google, San Francisco, USA
vrandecic@google.com

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**

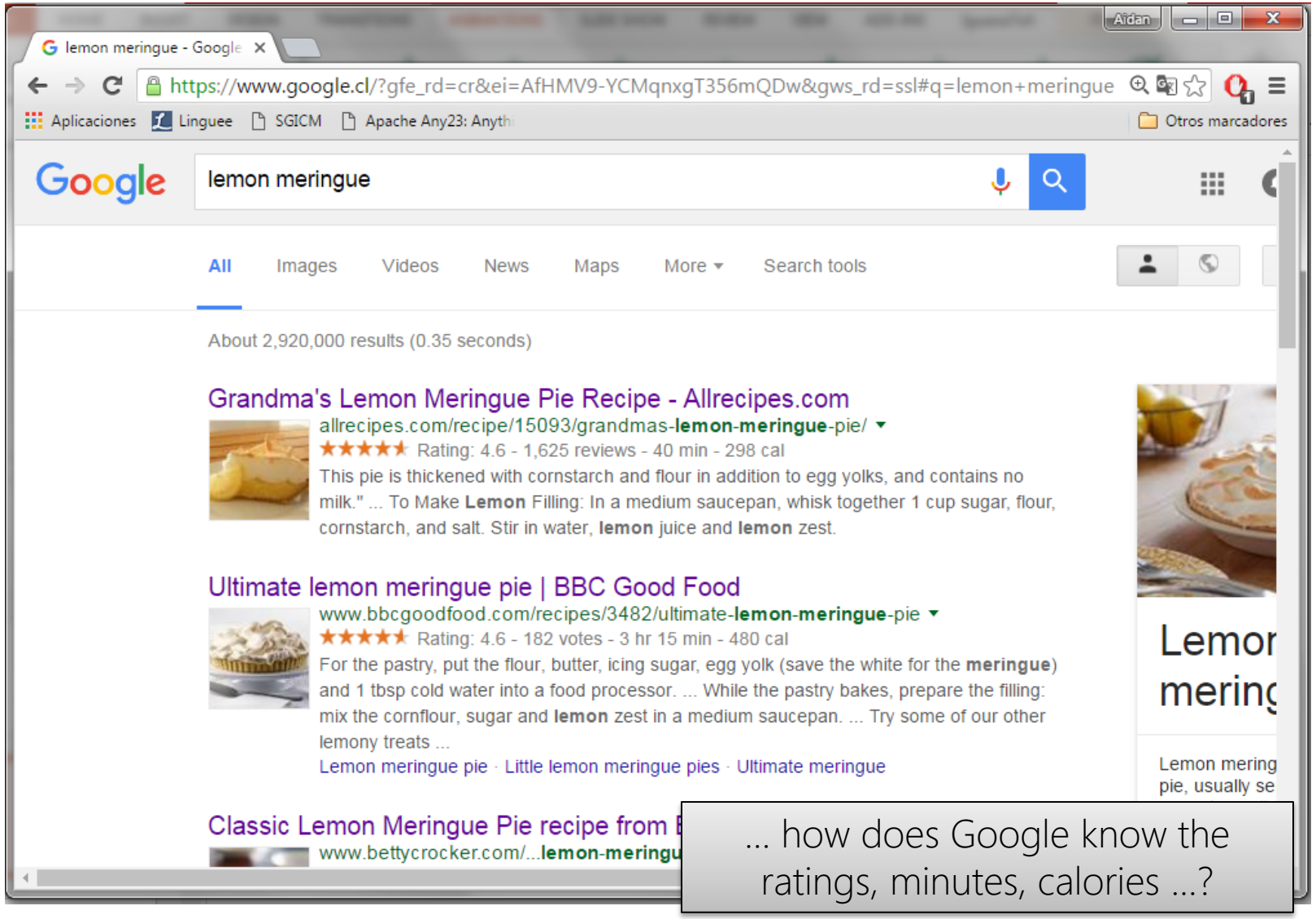
Collaborative knowledge bases that make their data freely available in a machine-readable form are central for the data strategy of many projects and organizations. The two major collaborative knowledge bases are Wikimedia's Wikidata and Google's Freebase. Due to the success of Wikidata, Google decided 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 …




… how does Google know the ratings, minutes, calories …?

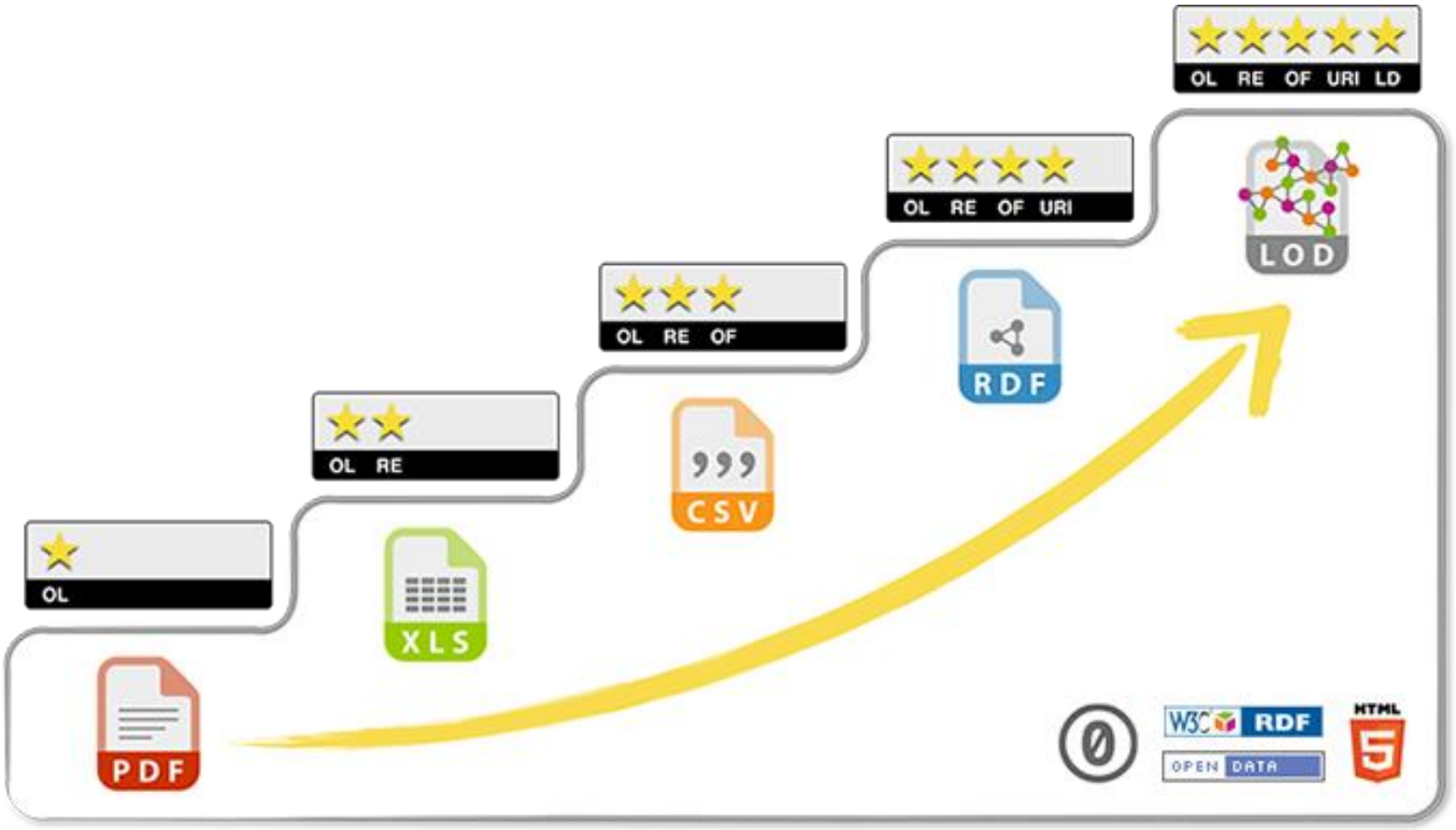
# Publishers add structured data!

```
<ul class="nutrition" itemprop="nutrition" itemscope
  itemtype="http://schema.org/NutritionInformation">
  <li>
    <span class="nutrition__label">kcal</span>
    <span class="nutrition__value" itemprop="calories">480</span>
    ...
  </li>
</ul>
```

- ✓ 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

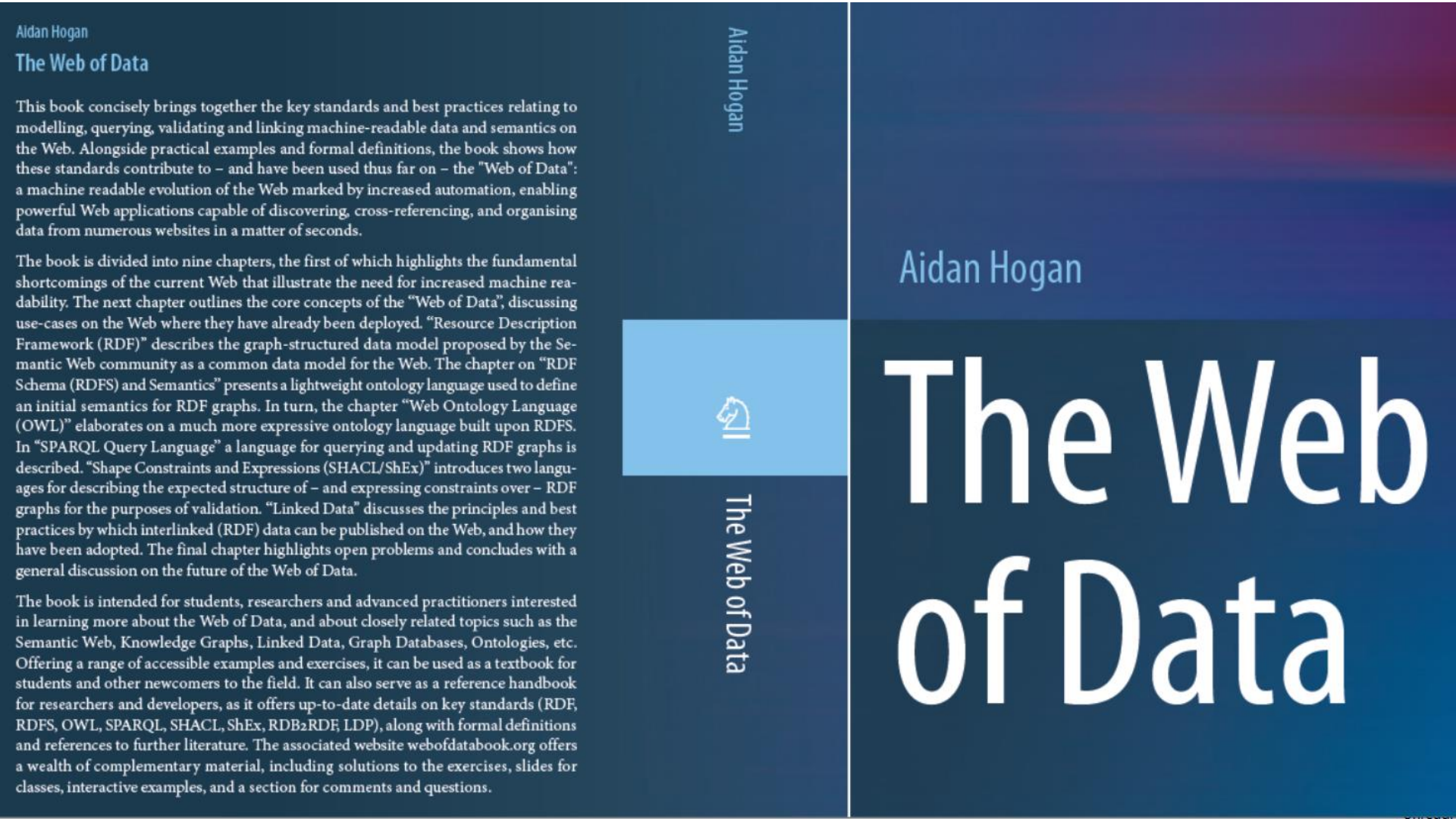
# Structure of the course

- 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



# Outcomes: Learn About the Semantic Web!

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

# Questions?