

CC3201-1

BASES DE DATOS

OTOÑO 2021

Clase 13: Datos Semiestructurados: Grafos

Aidan Hogan

aidhog@gmail.com

MODELOS DE DATOS

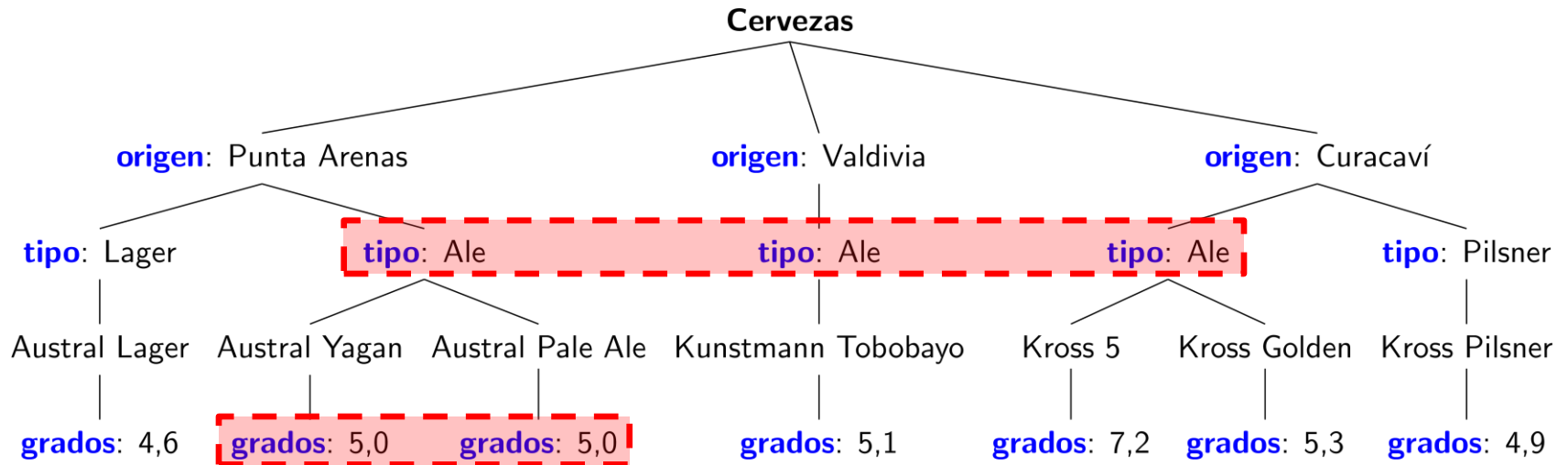
Modelo de datos (tabla)

Cervezas

nombre	tipo	grados	ciudad-origen
Austral Lager	Lager	4,6	Punta Arenas
Austral Yagan	Ale	5,0	Punta Arenas
Austral Pale Ale	Ale	5,0	Punta Arenas
Kuntsmann Torobayo	Ale	5,1	Valdivia
Kross 5	Ale	7,2	Curacaví
Kross Golden	Ale	5,3	Curacaví
Kross Pilsner	Pilsner	4,9	Curacaví



Modelo de datos (árbol/jerararquía)

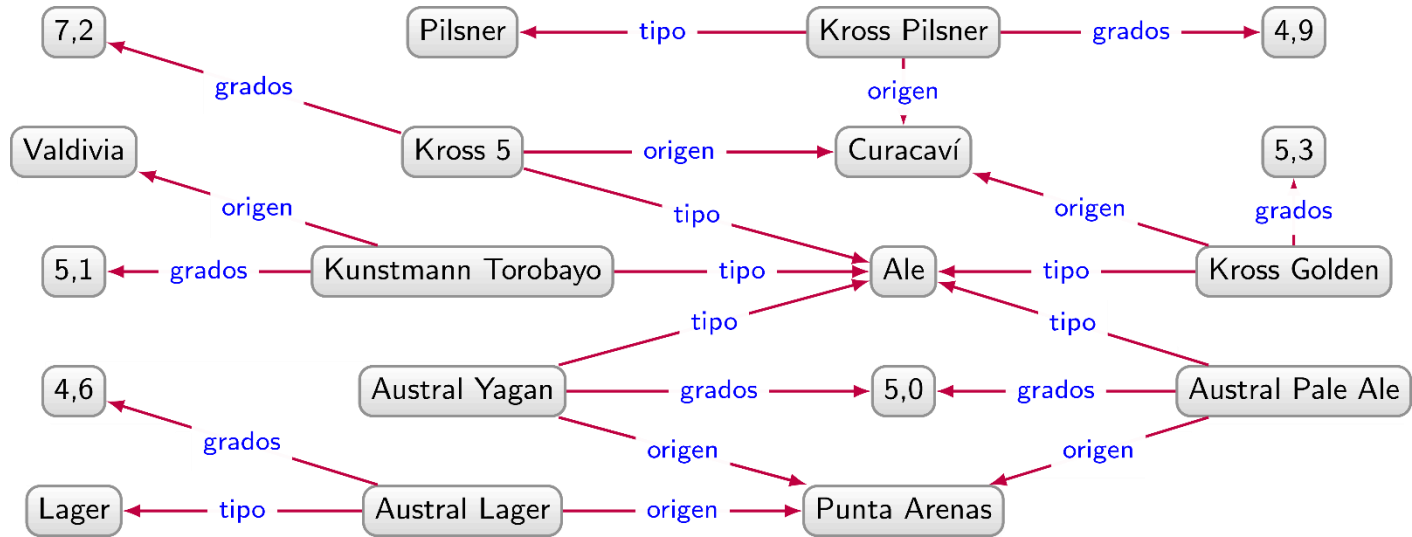


¿Es un “verdadero” árbol?

Umm ... sí y no.

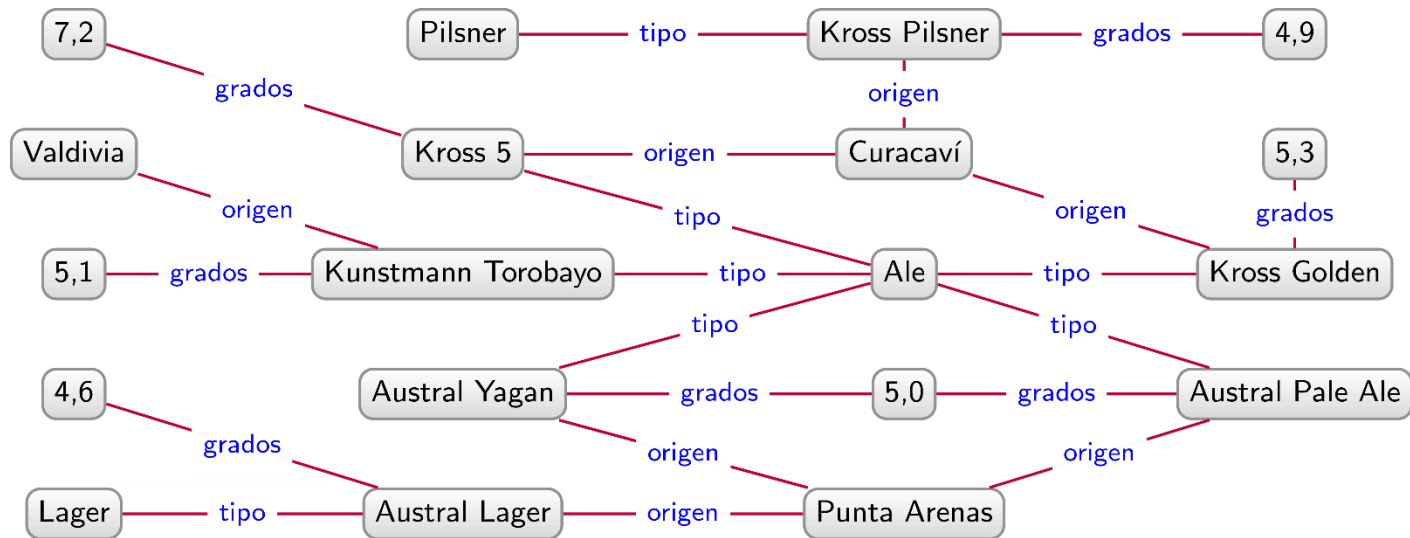
Es feo tener tres nodos para Ale
(y es más difícil de consultar)

Modelo de datos (grafo)



Un grafo *dirigido* con *arcos etiquetados*

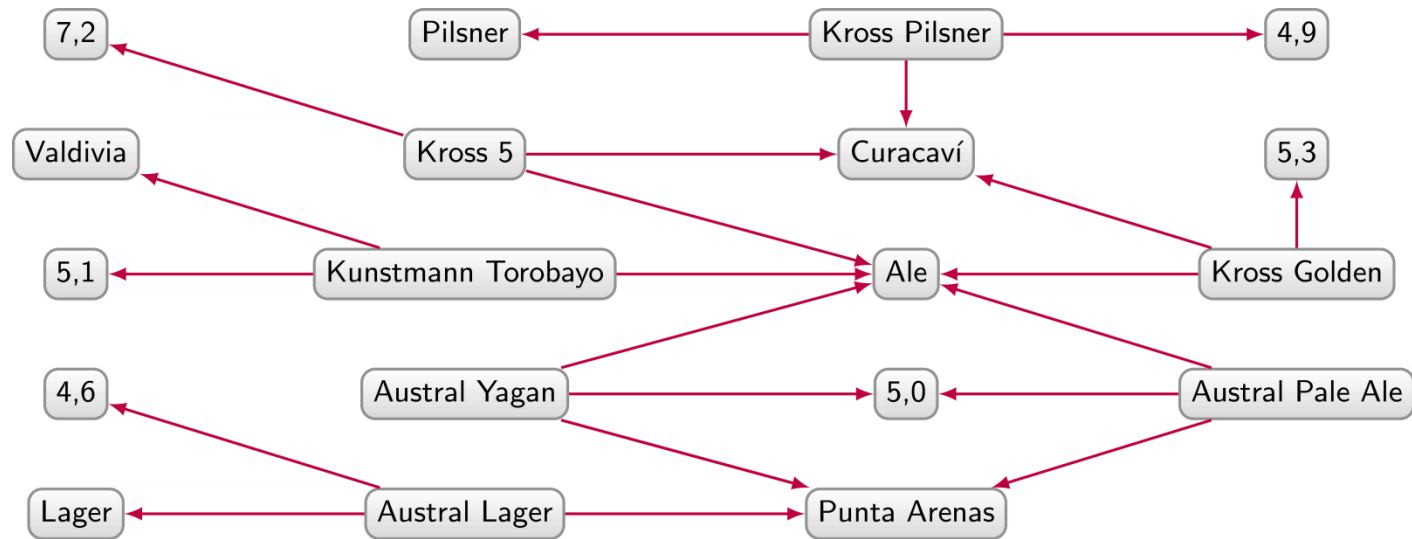
Grafo dirigido ...



¿El problema sin dirección?

¿Curacaví es el origen de Kross 5 o
Kross 5 es el origen de Curacaví?

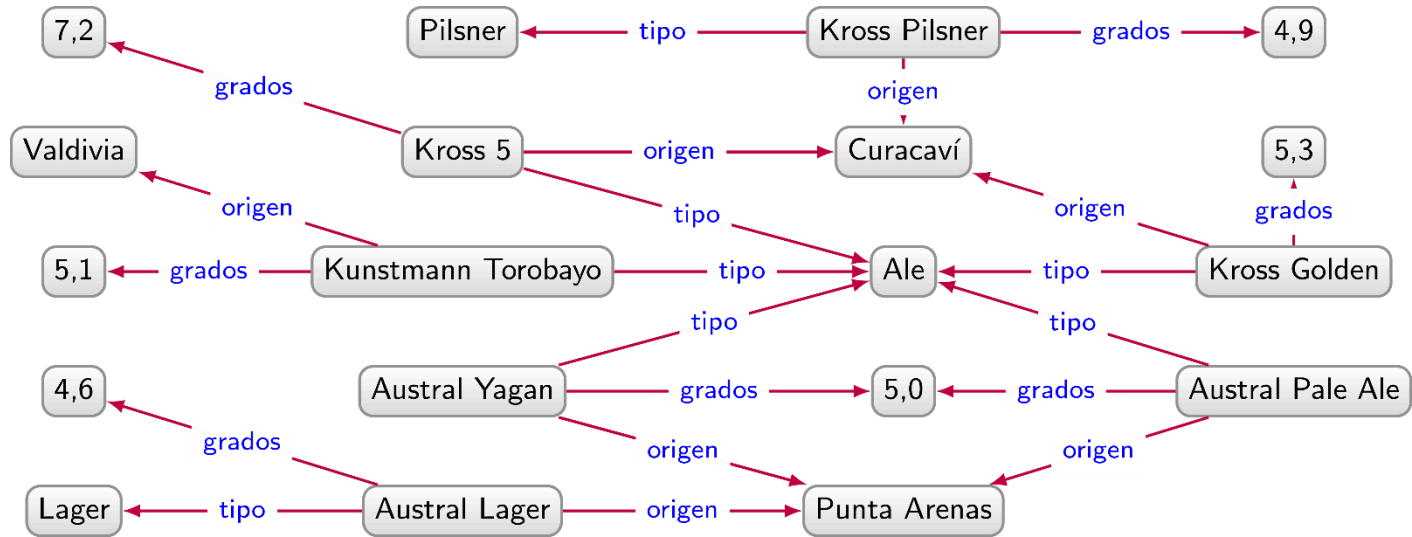
Grafo con arcos etiquetados



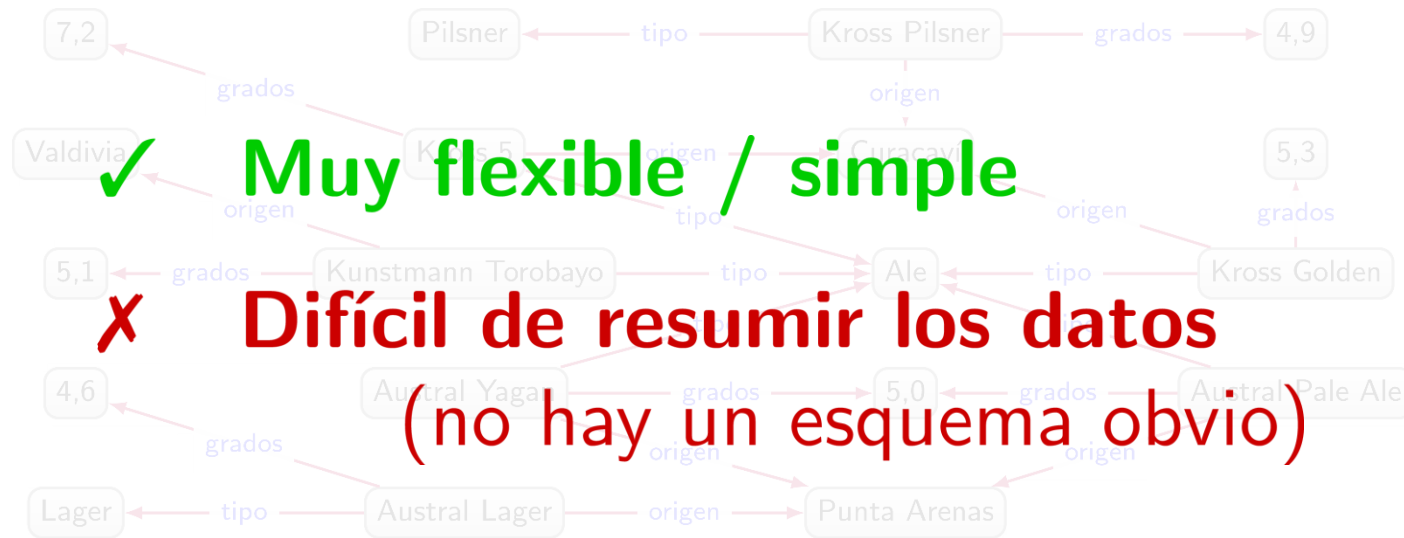
¿El problema sin etiquetas sobre los arcos?

No sabemos la relación entre Ale y Kross Golden, por ejemplo.

Un grafo dirigido con arcos etiquetados



Un grafo **dirigido** con **arcos etiquetados**



El espectro de la estructura de datos

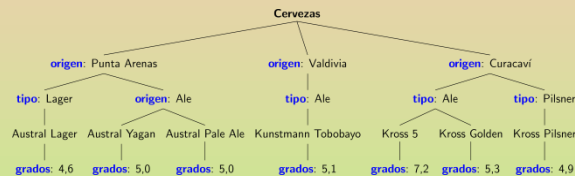
Texto Plano

Hay mucha variedad en las cervezas locales de Chile. La sede de la marca "Austral" se encuentra en Punta Arenas. Austral fabrica una amplia gama de cervezas, incluyendo Lager (4,6%), Yagan (un ale de 5%) y un Pale Ale (5%). La cerveza de marca "Kunstmann" también es popular, en particular su cerveza "Torobayo" (un ale de 5,1% elaborado en Valdivia). La marca Kross, basada en Curacaví, también tiene una gama de cervezas populares como, por ejemplo, Kross 5 (un ale fuerte de 7,2%) Kross Golden (un ale de 5,3%) y Kross Pilsner (4,9%).

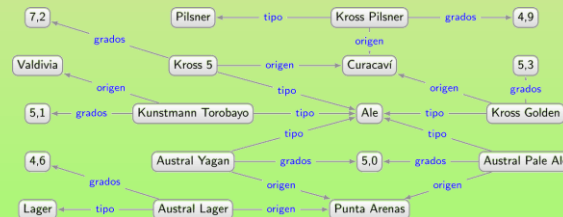
```
<ul>
<li>Austral Lager [Lager] 4,6% de Punta Arenas</li>
<li>Austral Yagan [Ale] 5% de Punta Arenas</li>
<li>Austral Pale Ale [Ale] 5% de Punta Arenas</li>
<li>Kunstmann Torobayo [Ale] 5,1% de Valdivia</li>
<li>Kross 5 [Ale] 7,2% de Curacaví</li>
<li>Kross Golden [Ale] 5,3% de Curacaví</li>
<li>Kross Pilsner [Pilsner] 4,9% de Curacaví</li>
</ul>
```

Texto Enriquecido
(HTML, Word, ...)

Árboles
(XML, JSON, ...)



Grafos
(RDF, Prop. Gs, ...)



Relacional
(SQL, CSV, ...)

Cervezas			
nombre	tipo	grados	ciudad-origen
Austral Lager	Lager	4,6	Punta Arenas
Austral Yagan	Ale	5,0	Punta Arenas
Austral Pale Ale	Ale	5,0	Punta Arenas
Kuntsmann Torobayo	Ale	5,1	Valdivia
Kross 5	Ale	7,2	Curacaví
Kross Golden	Ale	5,3	Curacaví
Kross Pilsner	Pilsner	4,9	Curacaví

No estructurados

D

A

T

O

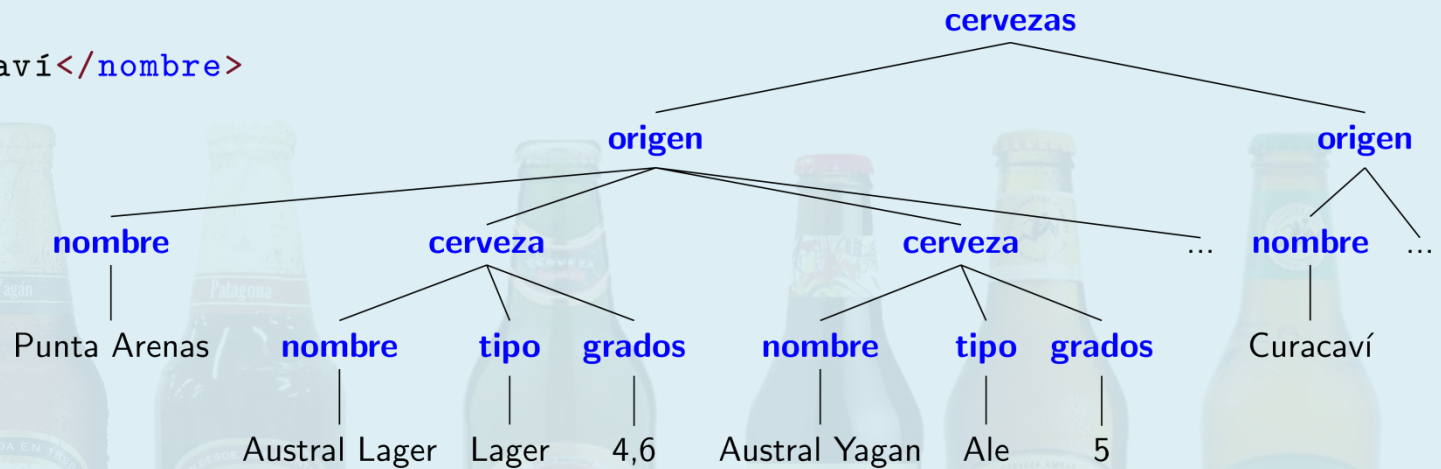
S

Semiestructurados

Estructurados

Para árboles, tenemos formatos como XML

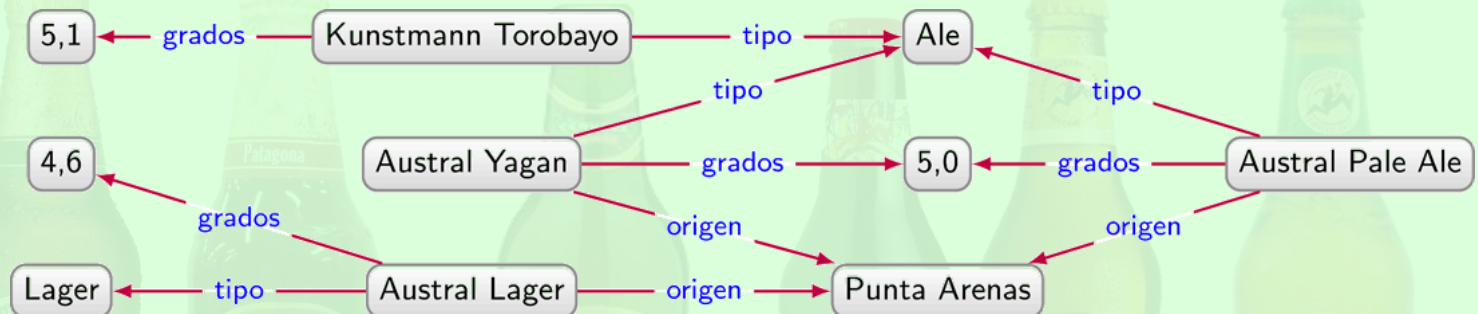
```
<cervezas>
  <origen>
    <nombre>Punta Arenas</nombre>
    <cerveza>
      <nombre>Austral Lager</nombre>
      <tipo>Lager</tipo>
      <grados>4.6</grados>
    </cerveza>
    <cerveza>
      <nombre>Austral Yagan</nombre>
      <tipo>Ale</tipo>
      <grados>5</grados>
    </cerveza>
    ...
  </origen>
  <origen>
    <nombre>Curacaví</nombre>
    ...
  </origen>
</cervezas>
```



Para grafos, podemos usar triples

Kunstmann_Torobayo	tipo	Ale
Kunstmann_Torobayo	grados	5,1
Austral_Yagan	tipo	Ale
Austral_Yagan	grados	5,0
Austral_Yagan	origen	Punta_Arenas
Austral_Pale_Ale	tipo	Ale
Austral_Pale_Ale	grados	5,0
Austral_Pale_Ale	origen	Punta_Arenas
Austral_Lager	tipo	Lager
Austral_Lager	grados	4,6
Austral_Lager	origen	Punta_Arenas

No hay orden.



JUEGO DE DATOS

Texto Plano

Texto Enriquecido
(HTML, Word, ...)

Árboles
(XML, JSON, ...)

Grafos
(RDF, Prop. Gs, ...)

Relacional
(SQL, CSV, ...)



D
A
T
O
S

No estructurados

Semiestructurados

Estructurados

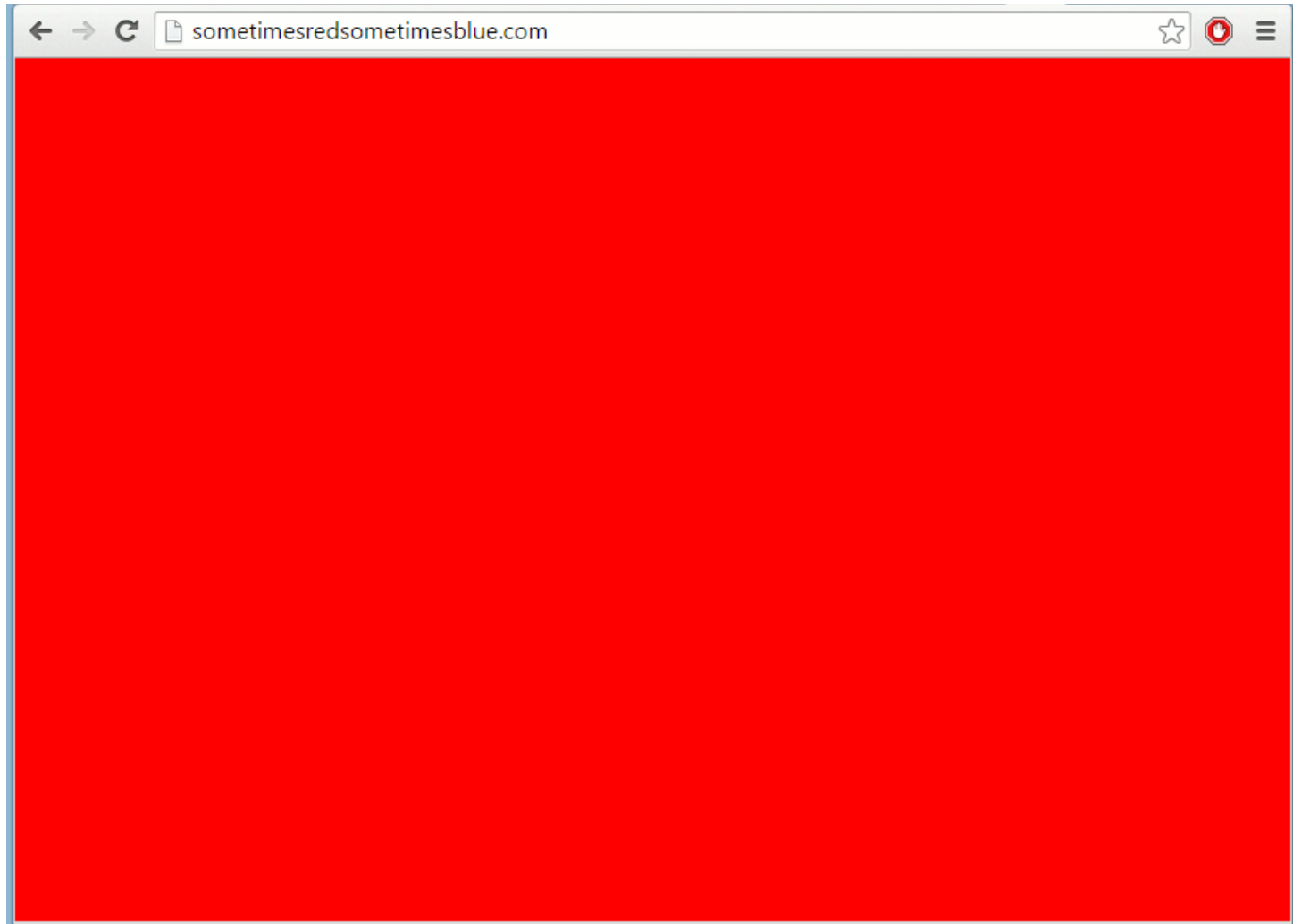
¿PERO DÓNDE DE USAN GRAFOS?

La Web de Datos / La Web Semántica



**Semantic
Web**

¡La Web es magnífica!



¿Pero puede ser mejor?

... a veces es difícil encontrar información relevante.

Haciendo un informe para su clase ...



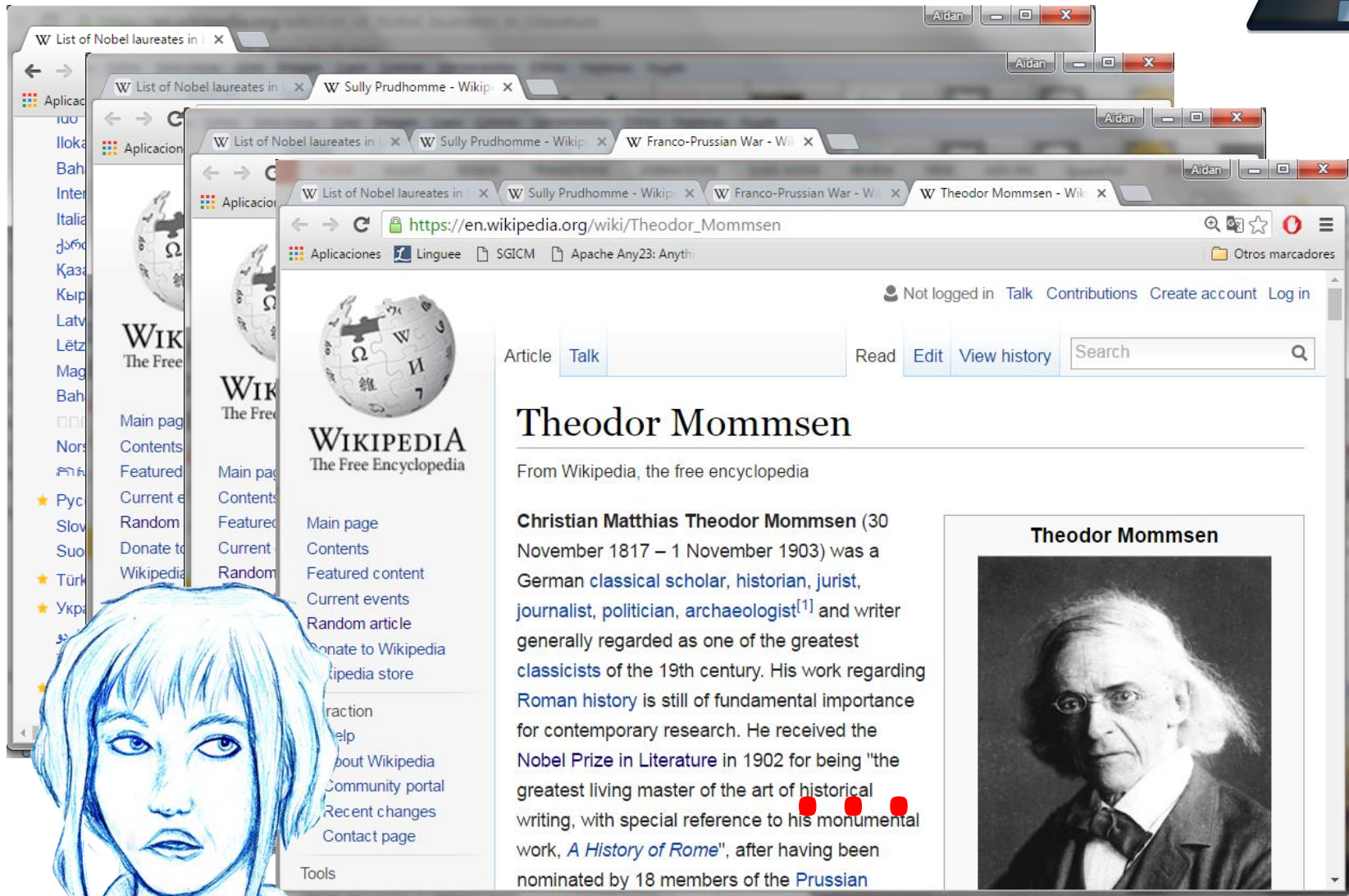
Quiere encontrar *cada*:

- Ganadores del Premio Nobel en Literatura
- Que han luchado en una guerra
- El año que ganaron el premio
- Y el año que comenzó la guerra



¿Cómo se puede encontrar esta información con la Web actual?

Muchas pestañas de Wikipedia ...



La Web en el espectro de datos

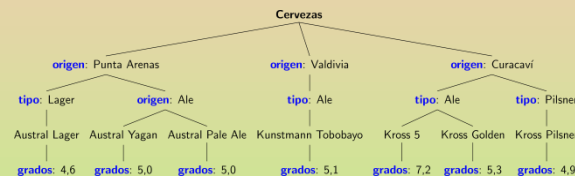
Texto Plano

Hay mucha variedad en las cervezas locales de Chile. La sede de la marca "Austral" se encuentra en Punta Arenas. Austral fabrica una amplia gama de cervezas, incluyendo Lager (4,6%), Yagan (un ale de 5%) y un Pale Ale (5%). La cerveza de marca "Kunstmann" también es popular, en particular su cerveza "Torobayo" (un ale de 5,1% elaborado en Valdivia). La marca Kross, basada en Curacaví, también tiene una gama de cervezas populares como, por ejemplo, Kross 5 (un ale fuerte de 7,2%) Kross Golden (un ale de 5,3%) y Kross Pilsner (4,9%).

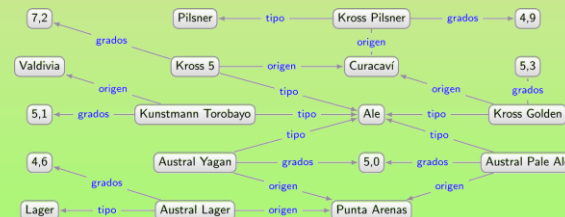
Texto Enrichido
(HTML, Word, ...)

```
<ul>
<li>Austral Lager [Lager] 4,6% de Punta Arenas</li>
<li>Austral Yagan [Ale] 5% de Punta Arenas</li>
<li>Austral Pale Ale [Ale] 5% de Punta Arenas</li>
<li>Kunstmann Torobayo [Ale] 5,1% de Valdivia</li>
<li>Kross 5 [Ale] 7,2% de Curacaví</li>
<li>Kross Golden [Ale] 5,3% de Curacaví</li>
<li>Kross Pilsner [Pilsner] 4,9% de Curacaví</li>
</ul>
```

Árboles
(XML, JSON, ...)



Grafos
(RDF, Prop. Gs, ...)



Relacional
(SQL, CSV, ...)

Cervezas			
nombre	tipo	grados	ciudad-origen
Austral Lager	Lager	4,6	Punta Arenas
Austral Yagan	Ale	5,0	Punta Arenas
Austral Pale Ale	Ale	5,0	Punta Arenas
Kuntsmann Torobayo	Ale	5,1	Valdivia
Kross 5	Ale	7,2	Curacaví
Kross Golden	Ale	5,3	Curacaví
Kross Pilsner	Pilsner	4,9	Curacaví

No estructurados

D
A
T
O
S

Semiestructurados

Estructurados

La Web Semántica ...



... una base de datos global usando la Web

La Web de Datos en el espectro de datos

Texto Plano

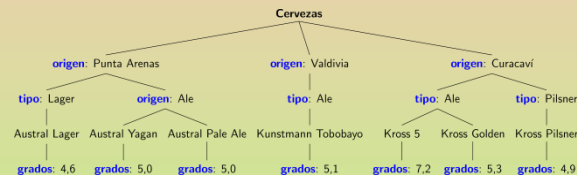
Hay mucha variedad en las cervezas locales de Chile. La sede de la marca "Austral" se encuentra en Punta Arenas. Austral fabrica una amplia gama de cervezas, incluyendo Larger (4,6%), Yagan (un ale de 5%) y un Pale Ale (5%). La cerveza de marca "Kunstmann" también es popular, en particular su cerveza "Torobayo" (un ale de 5,1% elaborado en Valdivia). La marca Kross, basada en Curacaví, también tiene una gama de cervezas populares como, por ejemplo, Kross 5 (un ale fuerte de 7,2%) Kross Golden (un ale de 5,3%) y Kross Pilsner (4,9%).

Texto Enriquecido (HTML, Word, ...)

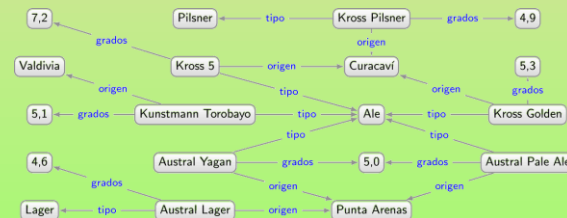
La Web

```
<ul>  
<li>Austral Lager [Lager] 4,6% de Punta Arenas</li>  
<li>Austral Lagan [Ale] 5% de Punta Arenas</li>  
<li>Austral Ale Ale [Ale] 4% de Punta Arenas</li>  
<li>Kunstsmann Torneo [Ale] 5% de Curacavi</li>  
<li>Kross 5% de Curacavi</li>  
<li>Kross Golden [Ale] 5,3% de Curacavi</li>  
<li>Kross Pilsner [Pilsner] 4,9% de Curacavi</li>  
</ul>
```

Árboles (XML, JSON, ...)



Grafos (RDF, Prop. Gs, ...)



Relacional
(SQL, CSV, ...)

Cervezas			
nombre	tipo	grados	ciudad-origen
Austral Lager	Lager	4,6	Punta Arenas
Austral Pilsner	Ale	5,1	Punta Arenas
Austral Pale Ale	Ale	5,1	Punta Arenas
Kentsmann Torpedo	Ale	5,1	Valdivia
Kross 5	Ale	7,2	Curacaví
Kross Golden	Ale	5,3	Curacaví
Kross Pilsner	Pilsner	4,9	Curacaví

No estructurados

¿Y si quisiéramos estructurar los datos de la Web y crear una Web de Datos, donde la pondríamos en el espectro?

Imaginemos el esquema relacional. :/

de Datos?

Estructurados

La Web de Datos en el espectro de datos

Texto Plano

Hay mucha variedad en las cervezas locales de Chile. La sede de la marca "Austral" se encuentra en Punta Arenas. Austral fabrica una amplia gama de cervezas, incluyendo Lager (4,6%), Yagan (un ale de 5%) y un Pale Ale (5%). La cerveza de marca "Kunstmann" también es popular, en particular su cerveza "Torobayo" (un ale de 5,1% elaborado en Valdivia). La marca Kross, basada en Curacaví, también tiene una gama de cervezas populares como, por ejemplo, Kross 5 (un ale fuerte de 7,2%) Kross Golden (un ale de 5,3%) y Kross Pilsner (4,9%).

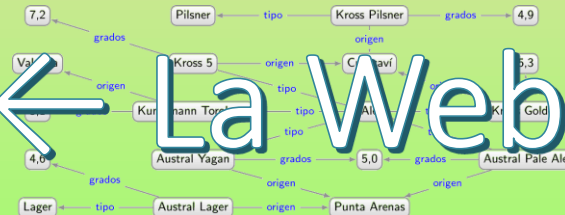
Texto Enrichido
(HTML, Word, ...)

```
<ul>
<li>Austral Lager [Lager] 4,6% de Punta Arenas</li>
<li>Austral Yagan [Ale] 5% de Punta Arenas</li>
<li>Austral Pale Ale [Ale] 5% de Punta Arenas</li>
<li>Kunstmann Torobayo [Ale] 5,1% de Valdivia</li>
<li>Kross 5 [Ale] 7,2% de Curacaví</li>
<li>Kross Golden [Ale] 5,3% de Curacaví</li>
<li>Kross Pilsner [Pilsner] 4,9% de Curacaví</li>
</ul>
```

Árboles
(XML, JSON, ...)



Grafos
(RDF, Prop. Gs, ...)



Relacional
(SQL, CSV, ...)

Cervezas			
nombre	tipo	grados	ciudad-origen
Austral Lager	Lager	4,6	Punta Arenas
Austral Yagan	Ale	5,0	Punta Arenas
Austral Pale Ale	Ale	5,0	Punta Arenas
Kuntsmann Torobayo	Ale	5,1	Valdivia
Kross 5	Ale	7,2	Curacaví
Kross Golden	Ale	5,3	Curacaví
Kross Pilsner	Pilsner	4,9	Curacaví

No estructurados

¿Y si quisiéramos estructurar los datos de la Web y crear una Web de Datos, donde la pondríamos en el espectro?

Parece más razonable usar grafos dado su flexibilidad

La Web de Datos

¿Pero qué formato de datos podemos ocupar entonces?

Estructura

...

Formato de Descripción de Recursos

RESOURCE DESCRIPTION FRAMEWORK (RDF)

Modelando el mundo con triples

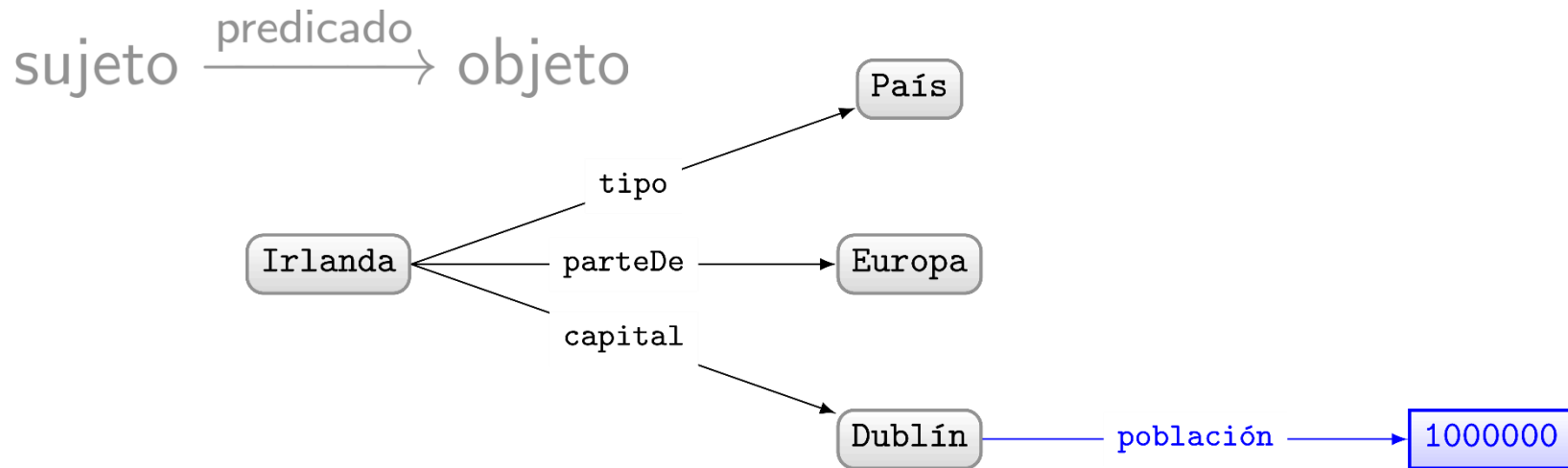
<i>sujeto</i>	<i>predicado</i>	<i>objeto</i>
Irlanda	parteDe	Europa
Irlanda	tipo	País
Irlanda	capital	Dublín

¿Agregar información? Concatenar triples.

<i>sujeto</i>	<i>predicado</i>	<i>objeto</i>
Irlanda	parteDe	Europa
Irlanda	tipo	País
Irlanda	capital	Dublín
Dublín	población	1,000,000

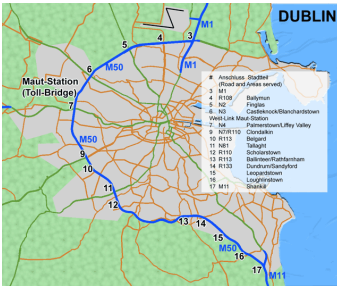
Los triples representan un grafo

<i>sujeto</i>	<i>predicado</i>	<i>objeto</i>
Irlanda	parteDe	Europa
Irlanda	tipo	País
Irlanda	capital	Dublín
Dublín	población	1,000,000



Los triples representan un grafo

<i>sujeto</i>	<i>predicado</i>	<i>objeto</i>
Irlanda	parteDe	Europa
Irlanda	tipo	País
	capital	Dublín
	población	1



sujeto

Irlanda

tipo
parteDe
capital

País

Europa

Dublín

población

1000000

Idea: Usar los identificadores de la Web

- **URL: *Uniform Resource Location***
 - La ubicación de un recurso de la Web
 - (p.ej., una página Web)
 - `http://ex.org/Dubl%C3%ADn.html`
- **URI: *Uniform Resource Identifier* (RDF 1.0)**
 - Un identificador de un recurso general
 - (p.ej., una ciudad)
 - `http://ex.org/Dubl%C3%ADn`
- **IRI: *Internationalised Resource Identifier* (RDF 1.1)**
 - Una URI que permite caracteres en *Unicode*
 - `http://ex.org/Dublín`

Usaremos IRIs con prefijos

`http://ex.org/Dublín` ↔ `ex:Dublín`

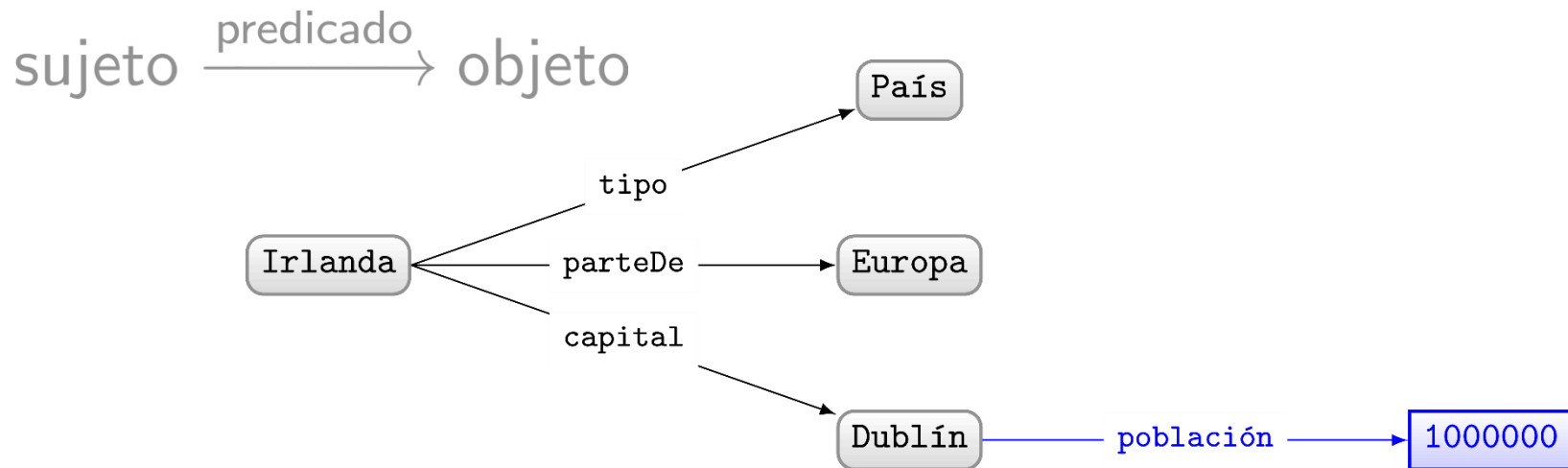
- “`ex:`” denota un prefijo para `http://ex.org/`
- “`Dublín`” en un nombre local

Prefijos comunes:

Prefix	Value
<code>rdf:</code>	<code>http://www.w3.org/1999/02/22-rdf-syntax-ns#</code>
<code>xsd:</code>	<code>http://www.w3.org/2001/XMLSchema#</code>
<code>rdfs:</code>	<code>http://www.w3.org/2000/01/rdf-schema#</code>
<code>owl:</code>	<code>http://www.w3.org/2002/07/owl#</code>

Desde strings ...

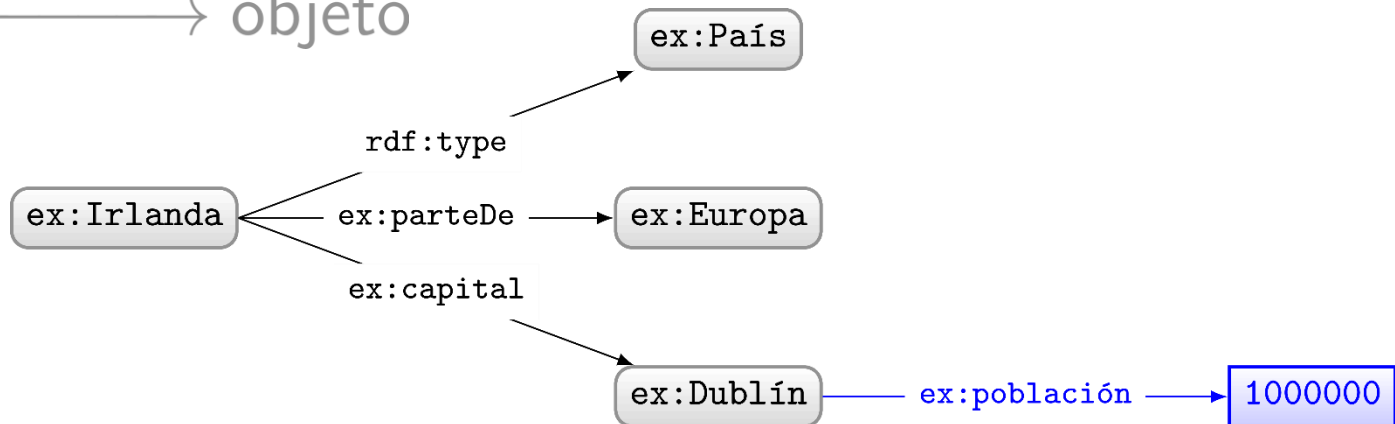
<i>sujeto</i>	<i>predicado</i>	<i>objeto</i>
Irlanda	parteDe	Europa
Irlanda	tipo	País
Irlanda	capital	Dublín
Dublín	población	1,000,000



... hasta identificadores de la Web ...

<i>sujeto</i>	<i>predicado</i>	<i>objeto</i>
ex:Irlanda	ex:parteDe	ex:Europa
ex:Irlanda	rdf:type	ex:País
ex:Irlanda	ex:capital	ex:Dublín
ex:Dublín	ex:población	1,000,000

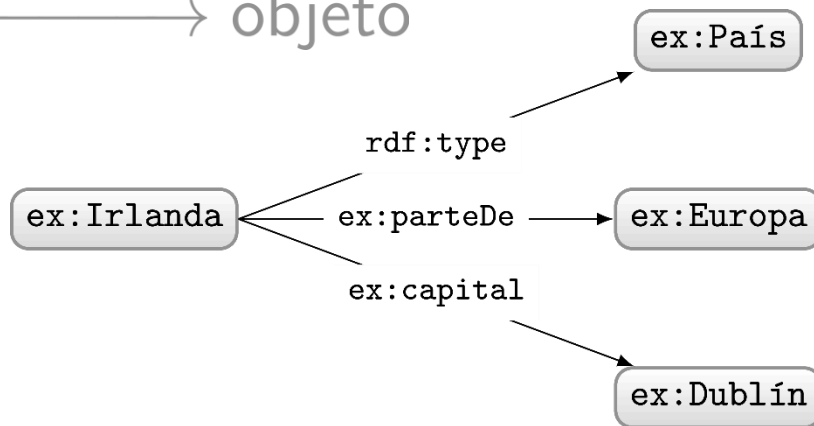
sujeto $\xrightarrow{\text{predicado}}$ objeto



... ¿y qué pasa con números, etc.? ...

<i>sujeto</i>	<i>predicado</i>	<i>objeto</i>
ex:Irlanda	ex:parteDe	ex:Europa
ex:Irlanda	rdf:type	ex:País
ex:Irlanda	ex:capital	ex:Dublín
ex:Dublín	ex:población	1,000,000

sujeto $\xrightarrow{\text{predicado}}$ objeto



¿Deberíamos asignar IRIs a números, booleanos, etc.?

Literals representan “valores” ...

... como strings, números, booleanos, fechas ...

- Pero solo se permiten *literals* como objetos

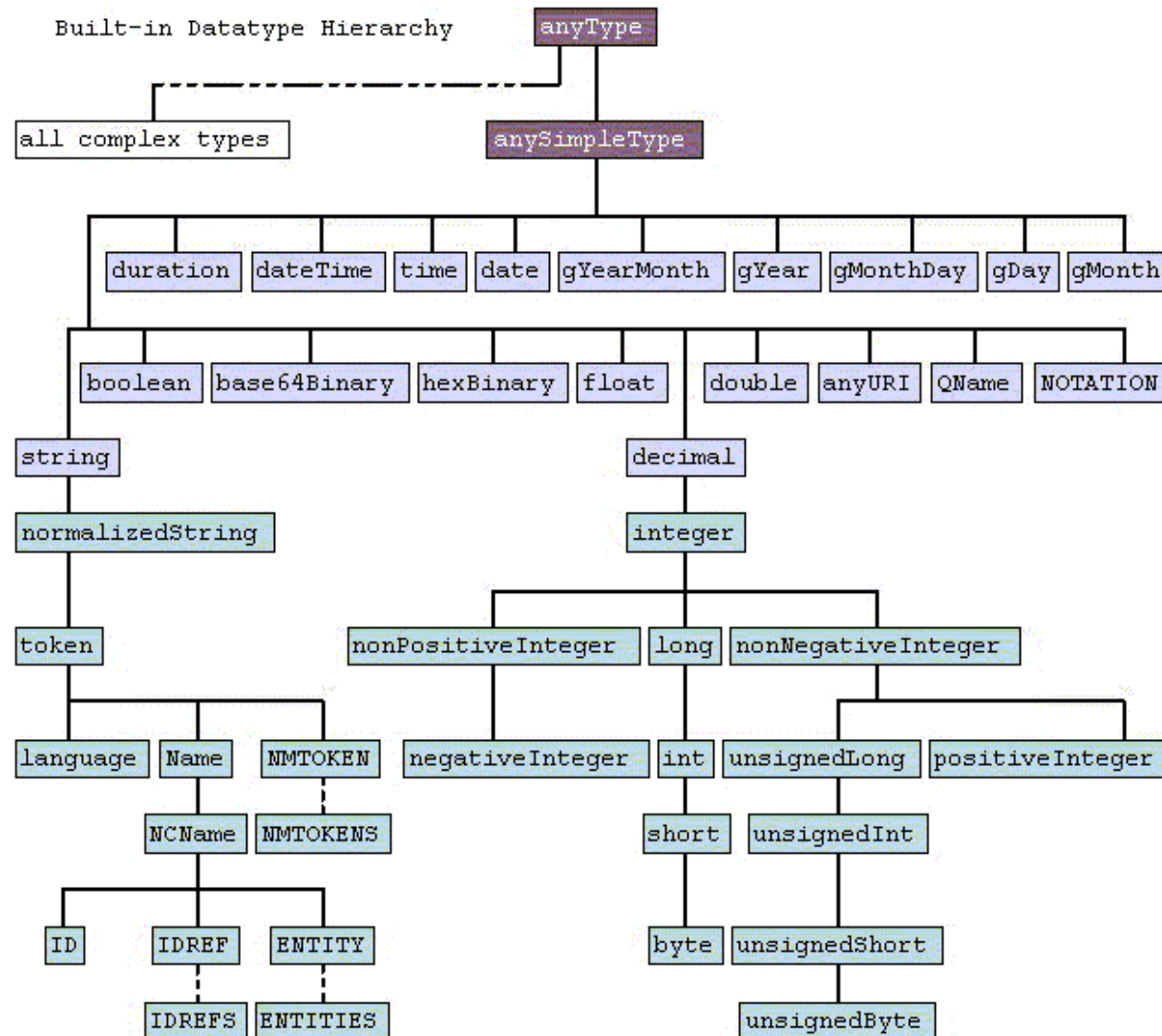
<i>sujeto</i>	<i>predicado</i>	<i>objeto</i>
ex:Dublín	ex:población	1,000,000
1,000,000	ex:población	ex:Dublín
ex:Dublín	1,000,000	ex:población

✓ CORRECT

✗ INCORRECT

✗ INCORRECT

Se usan los tipos de XML Schema



Se usan los tipos de XML Schema

"200"^^xsd:int

"true"^^xsd:boolean

"4.065"^^xsd:decimal

"2020-02-20"^^xsd:date

"2020-02-20T12:41:10Z"^^xsd:dateTimeStamp

"Sharknado"
Por defecto un string

Se pueden usar etiquetas de idioma

"200"^^xsd:int

"true"^^xsd:boolean

"4.065"^^xsd:decimal

"2020-02-20"^^xsd:date

"2020-02-20T12:41:10Z"^^xsd:dateTimeStamp

"Sharknado"

"Dublin"@en

"Dublín"@es-CL

"Baile Átha Cliath"@ga

Etiquetas de idioma

...

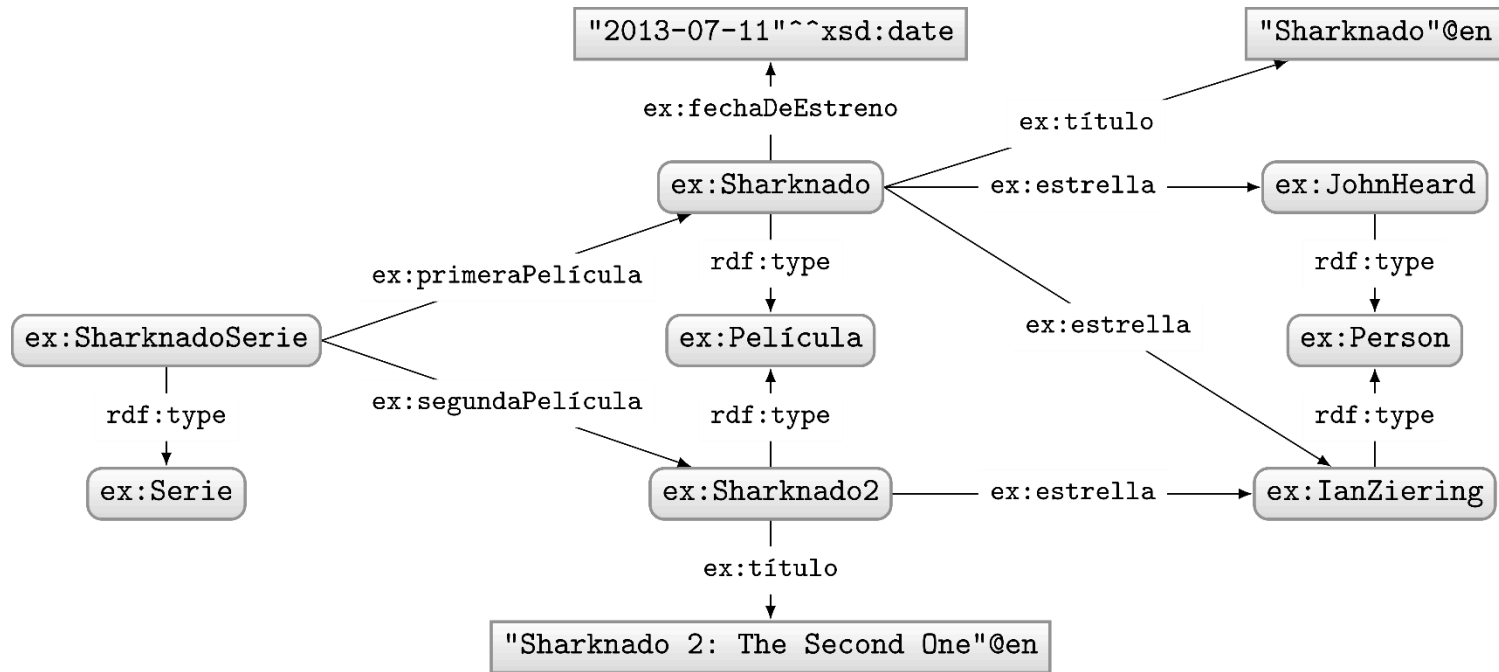
Protocolo y Lenguaje de Consulta de RDF:

SPARQL: SPARQL PROTOCOL AND RDF QUERY LANGUAGE



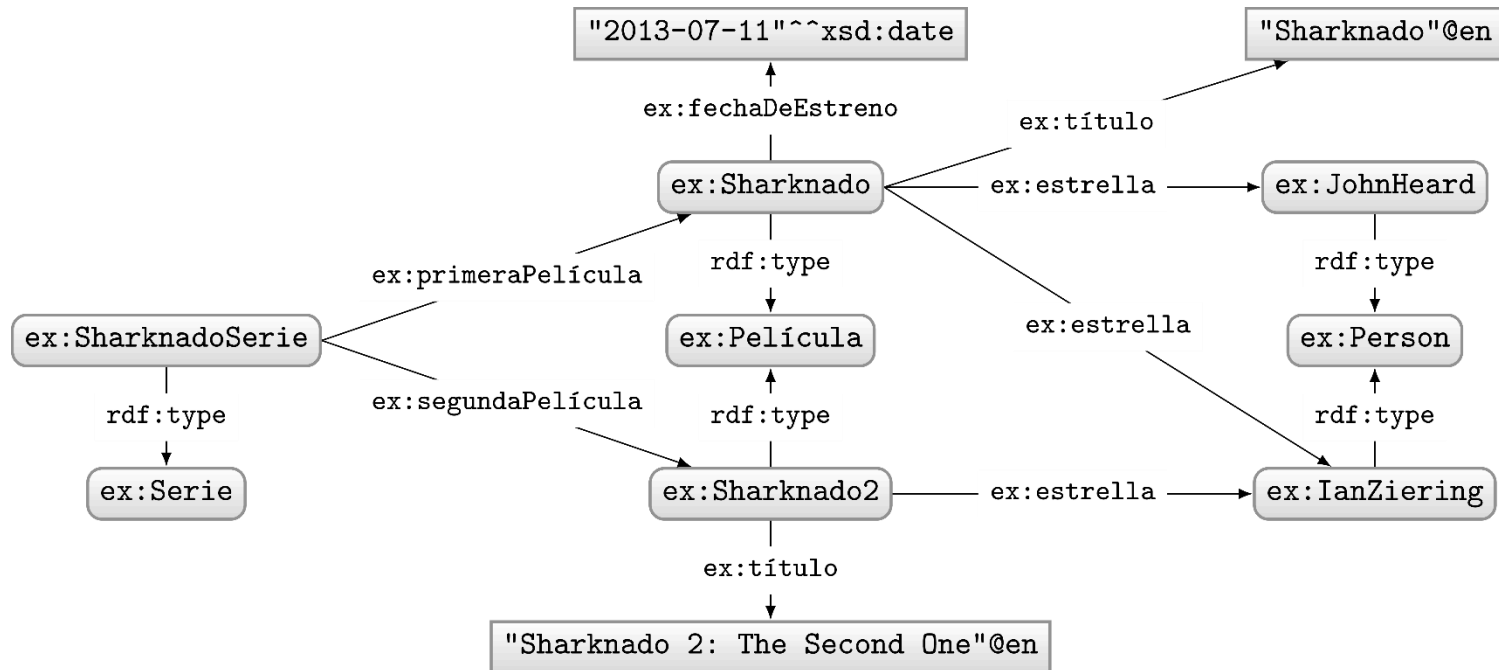
QOL (1.1)

SPARQL: Consultar Grafos en RDF



Consulta: “¿Quién protagoniza en la película ‘Sharknado’?”

SPARQL: Consultar Grafos en RDF



Consulta:

```
PREFIX ex: <http://ex.org/voc#>
SELECT *
WHERE {
  ex:Sharknado ex:estrella ?e .
}
```

Soluciones:

?e

ex:JohnHeard
ex:IanZiering

SPARQL Prefijos: Abreviaturas de IRIs

```
PREFIX ex: <http://ex.org/voc#>
SELECT *
WHERE {
    ex:Sharknado ex:estrella ?e .
}
```

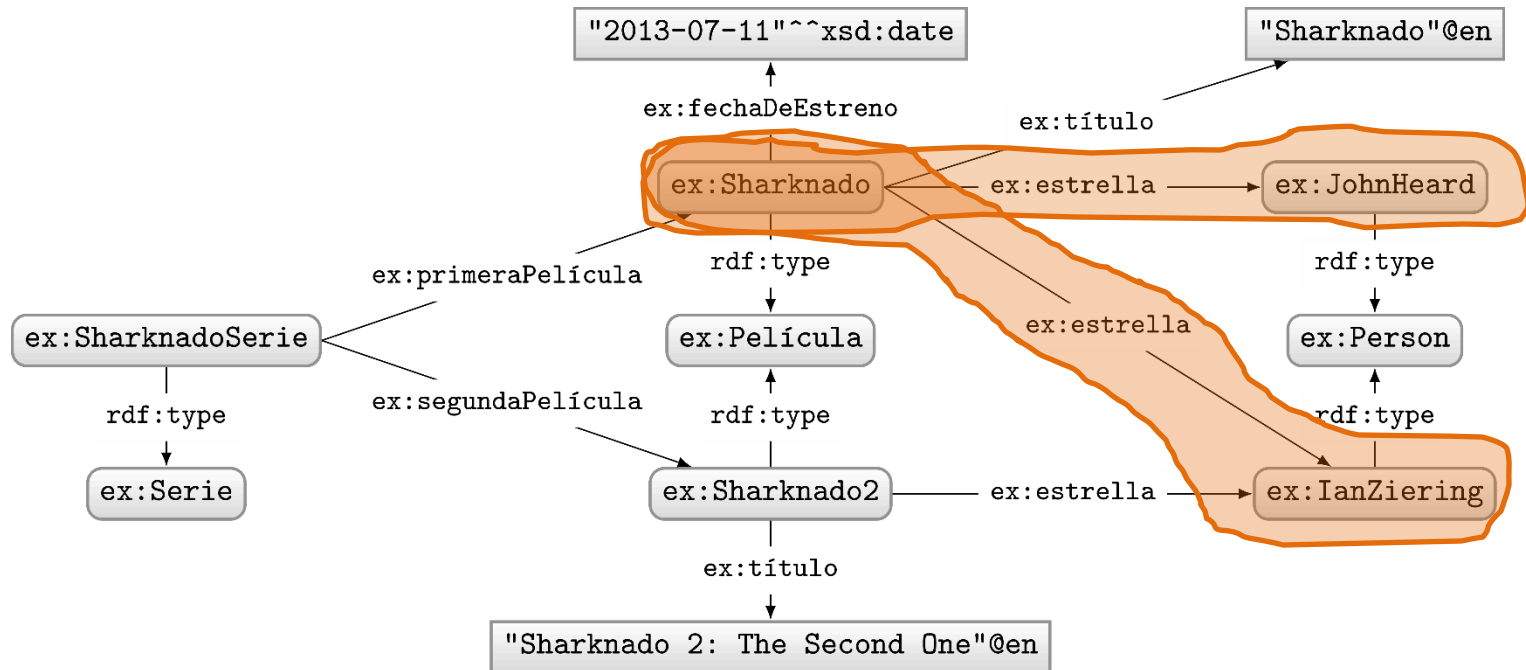
SPARQL: clausula de WHERE

- Especifica un grafo de consulta

```
PREFIX ex: <http://ex.org/voc#>
SELECT *
WHERE {
  ex:Sharknado ex:estrella ?e .
}
```

“Patrón triple”
(un triple con variables)

SPARQL: clausula de WHERE

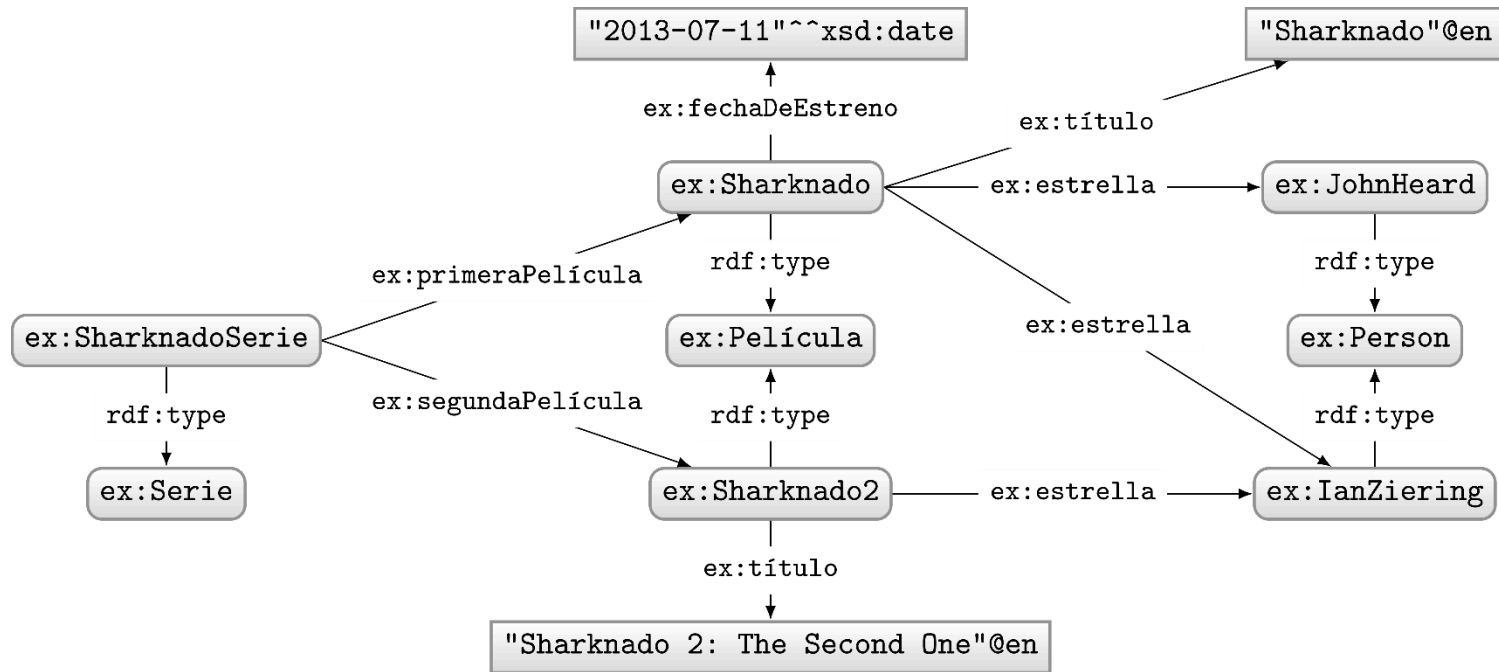


```
PREFIX ex: <http://ex.org/voc#>
SELECT *
WHERE {
  ex:Sharknado ex:estrella ?e .
}
```

?e

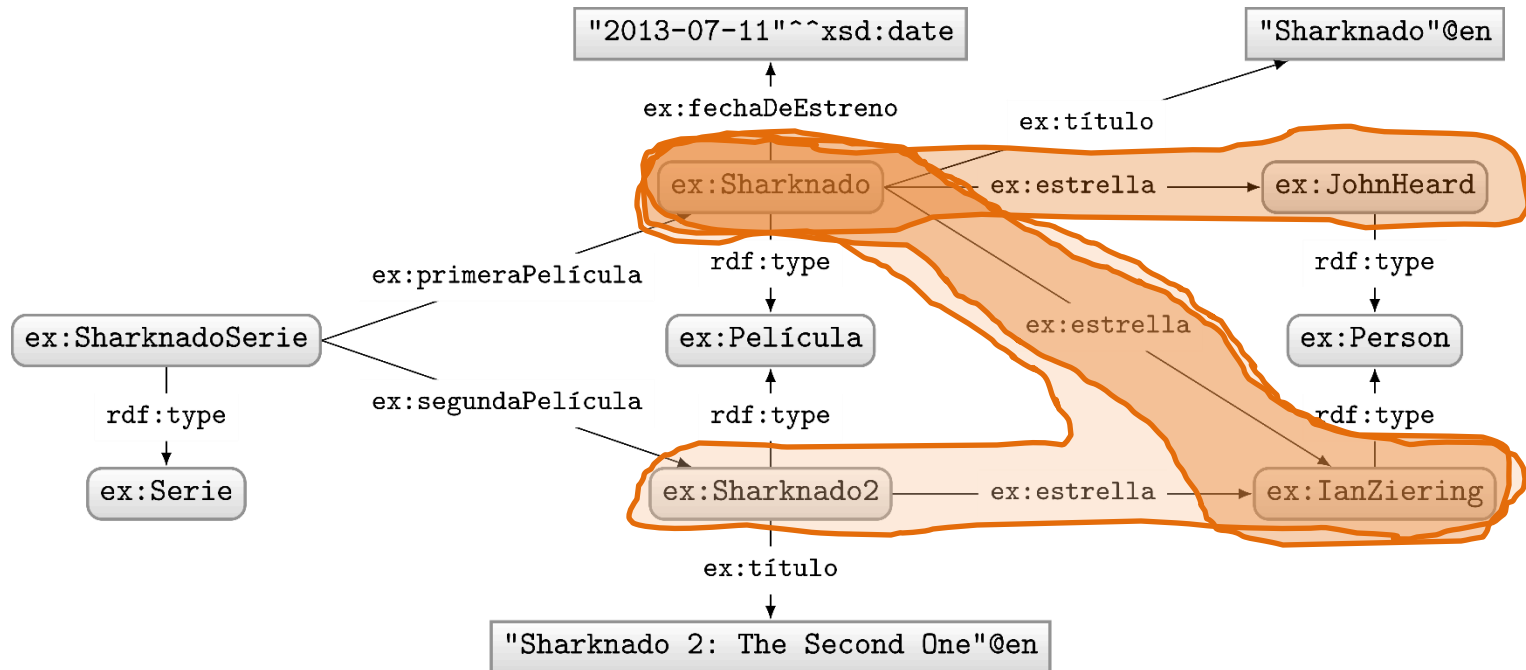
ex:JohnHeard
ex:IanZiering

SPARQL: clausula de WHERE



Consulta: “¿En cuáles (otras) películas han actuado los actores de Sharknado?”

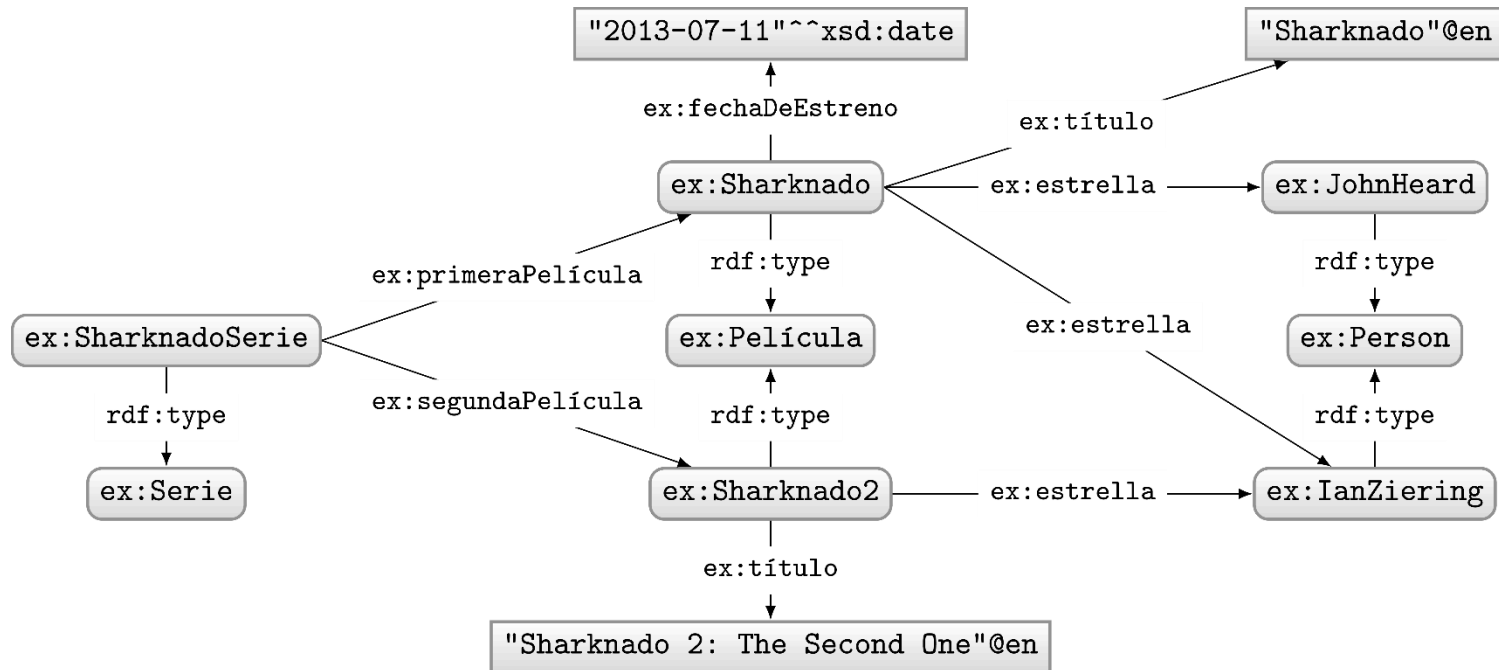
SPARQL: clausula de WHERE



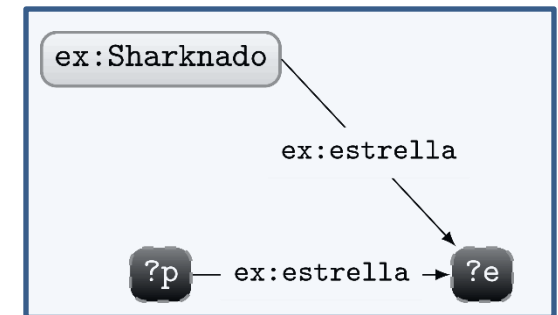
```
PREFIX ex: <http://ex.org/voc#>
SELECT *
WHERE {
  ex:Sharknado ex:estrella ?e .
  ?p ex:estrella ?e .
}
```

?e	?p
ex:IanZiering	ex:Sharknado2
ex:IanZiering	ex:Sharknado
ex:JohnHeard	ex:Sharknado

SPARQL: clausula de WHERE



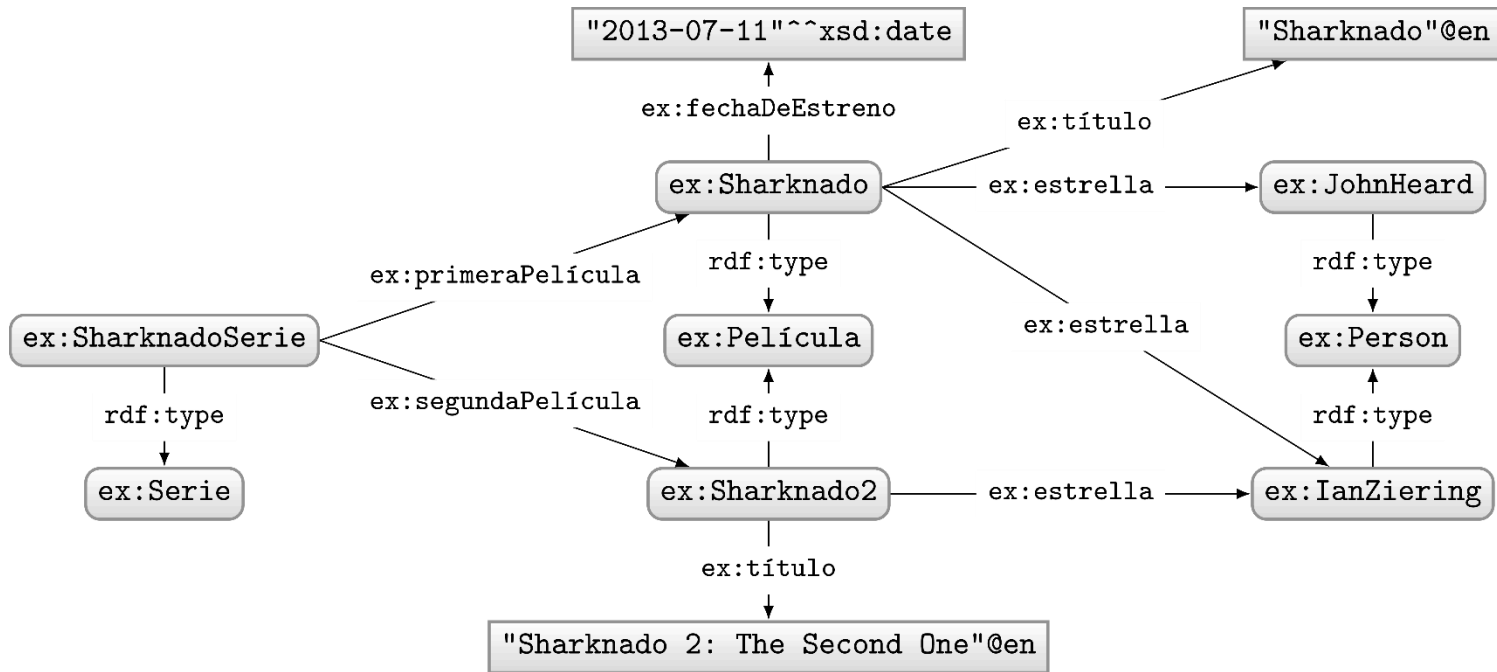
```
PREFIX ex: <http://ex.org/voc#>
SELECT *
WHERE {
  ex:Sharknado ex:estrella ?e .
  ?p ex:estrella ?e .
}
```



"Basic Graph Pattern"

(un conjunto de patrones triples)

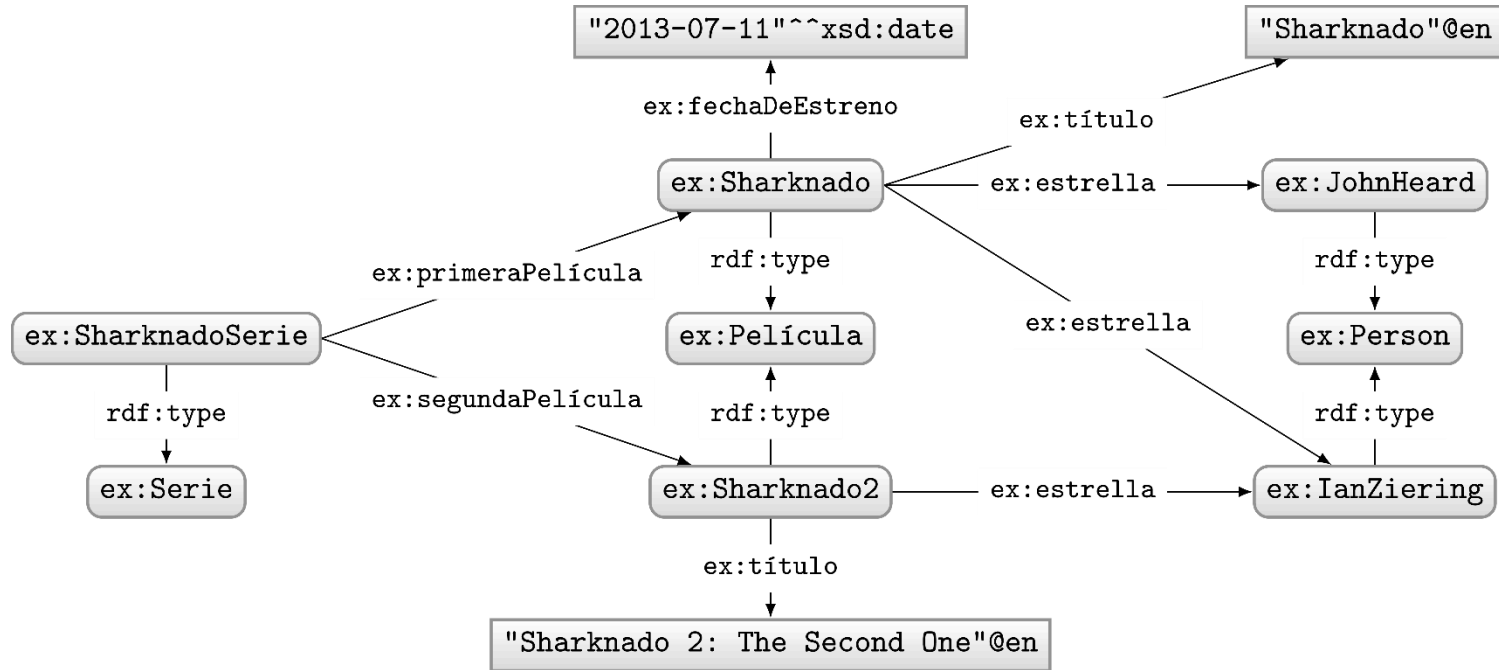
SPARQL: Joins



```
PREFIX ex: <http://ex.org/voc#>
SELECT *
WHERE {
    ex:Sharknado ex:estrella ?e.
    ?p ex:estrella ?e.
}
```

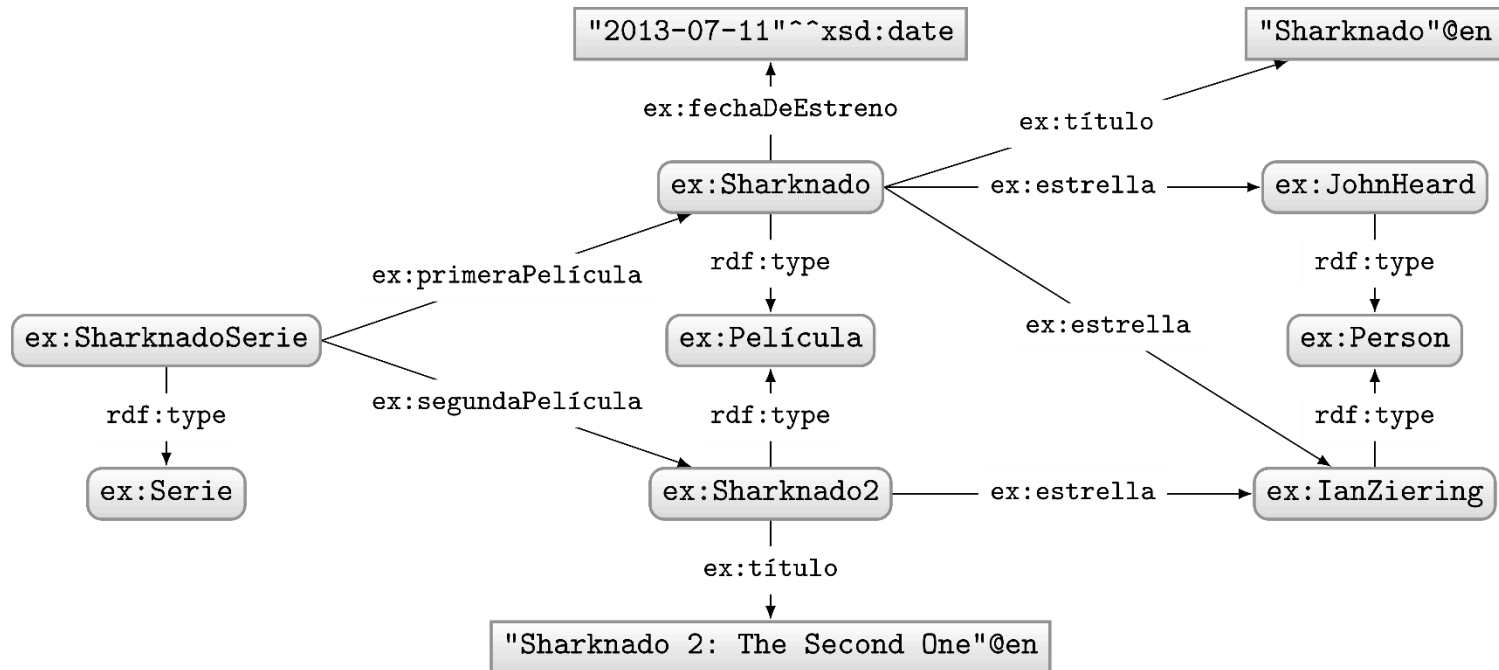
“Variable de Join”
(una variable en múltiples lugares)

SPARQL: Unión



Consulta: “¿Qué son los títulos de las dos primeras películas en la serie Sharknado?”

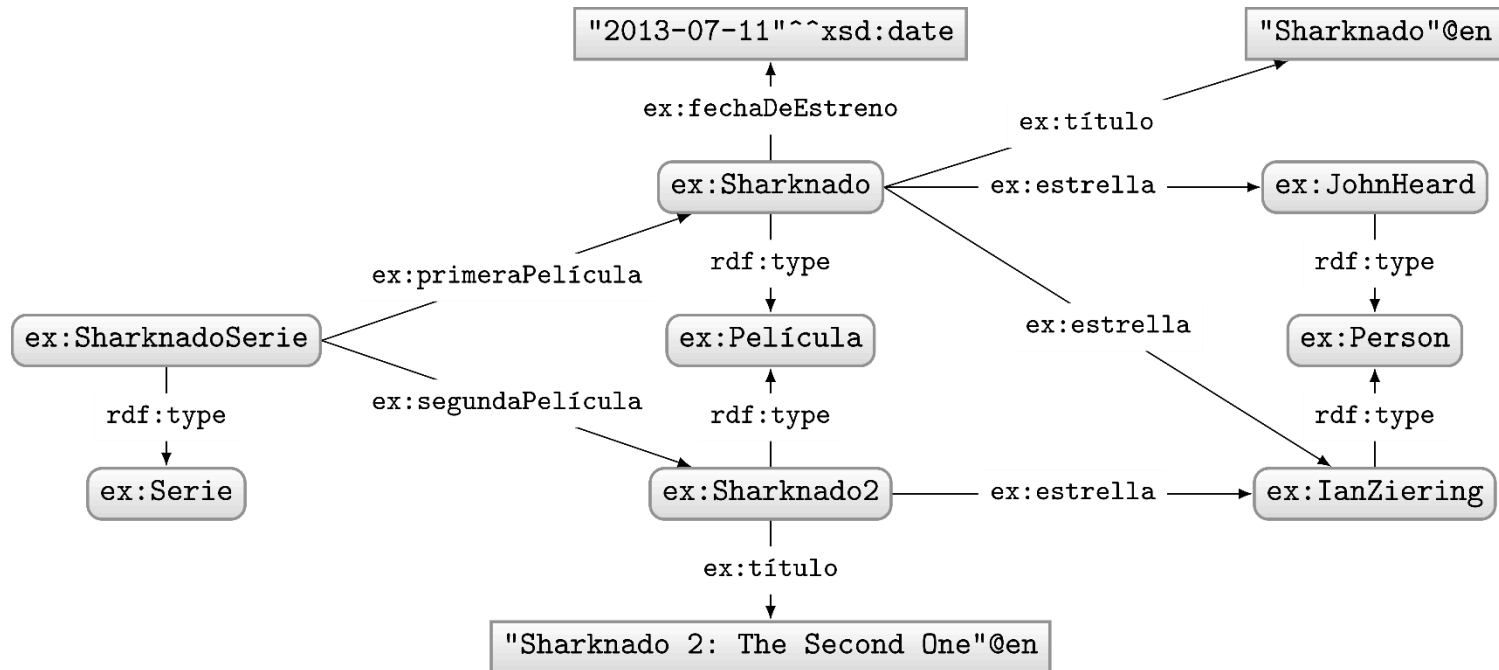
SPARQL: Unión



```
PREFIX ex: <http://ex.org/voc#>
SELECT *
WHERE {
  { ex:SharknadoSerie ex:primeraPelícula ?p . }
  UNION
  { ex:SharknadoSerie ex:segundaPelícula ?p . }
  ?p ex:título ?t .
}
```

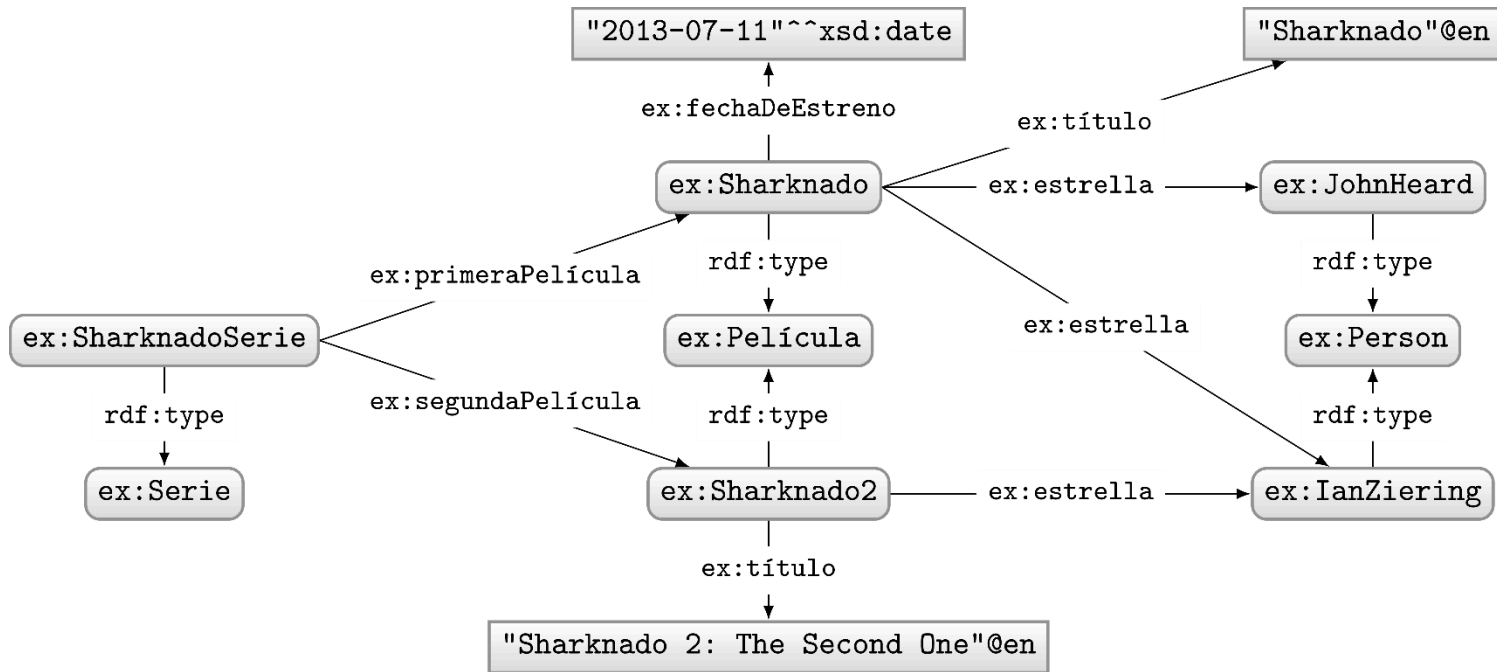
?p	?t
ex:Sharknado	"Sharknado"@en
ex:Sharknado2	"Sharknado 2: The Second One"@en

SPARQL: *Left-join* (OPTIONAL)



Consulta: “¿Los títulos de películas y (de estar disponible) sus fechas de estreno?”

SPARQL: *Left-join* (OPTIONAL)

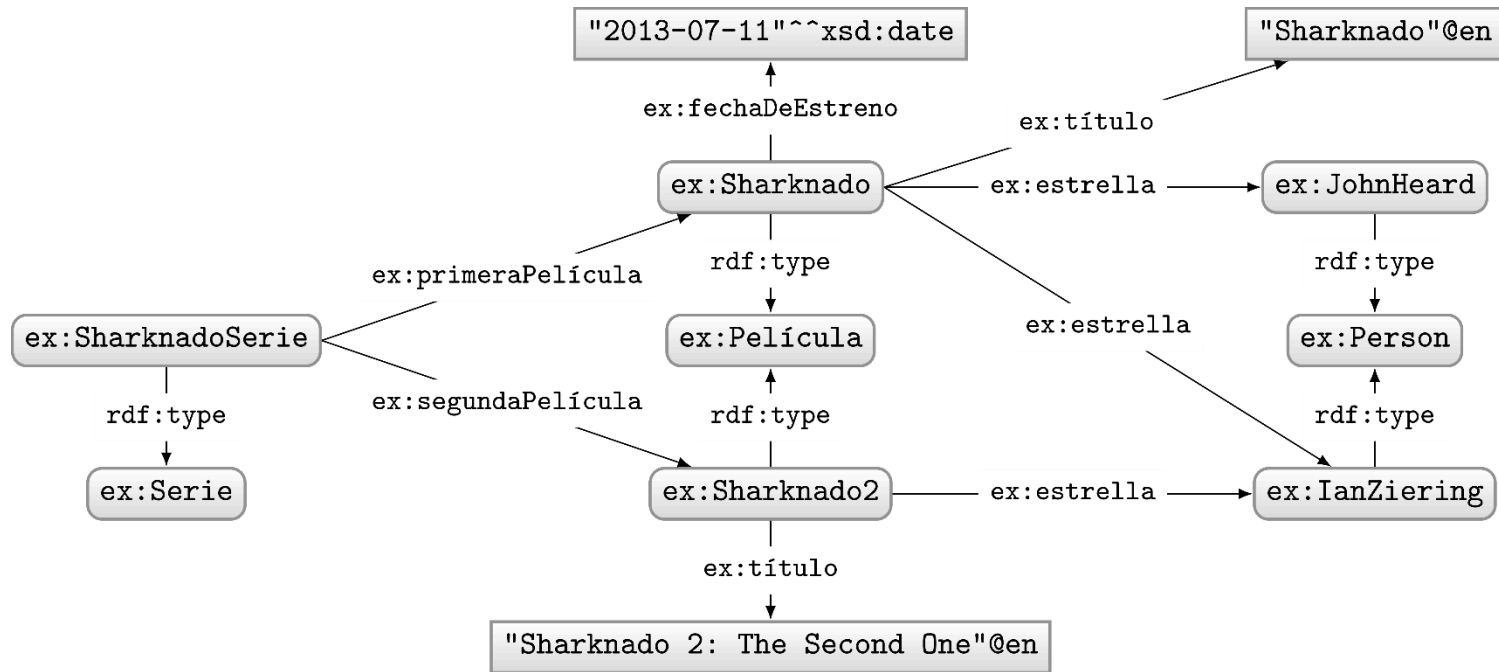


```
PREFIX ex: <http://ex.org/voc#>
SELECT *
WHERE {
    ?p a ex:Película .
    ?p ex:título ?t .
    OPTIONAL
        { ?p ex:fechaDeEstreno ?f }
}
```

?p	?t	?f
ex:Sharknado ex:Sharknado2	"Sharknado"@en "Sharknado 2: The Second One"@en	"2013-07-11"^^xsd:date 2013-07-11

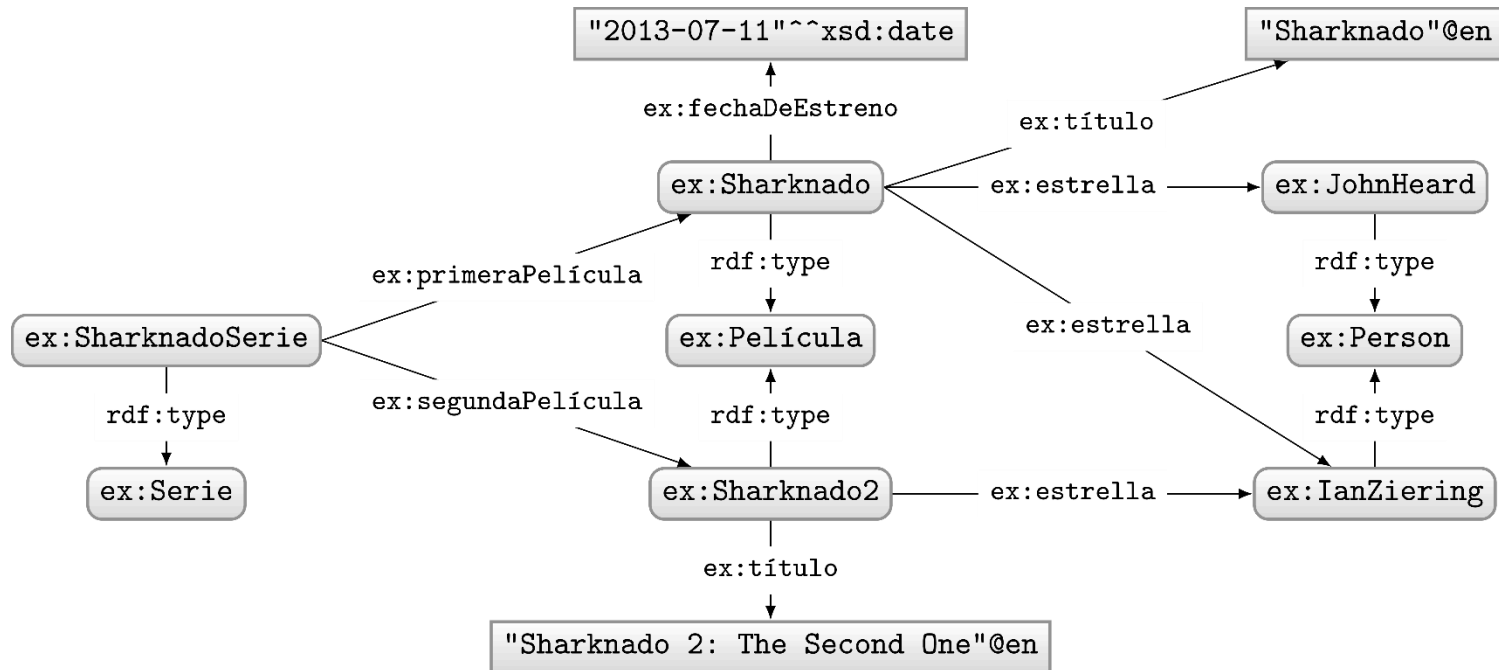
“Variable UNBOUND”
(una variable sin una solución)

SPARQL: Filtros



Consulta: “¿Cuáles películas estrenaron en 2014?”

SPARQL: Filtros

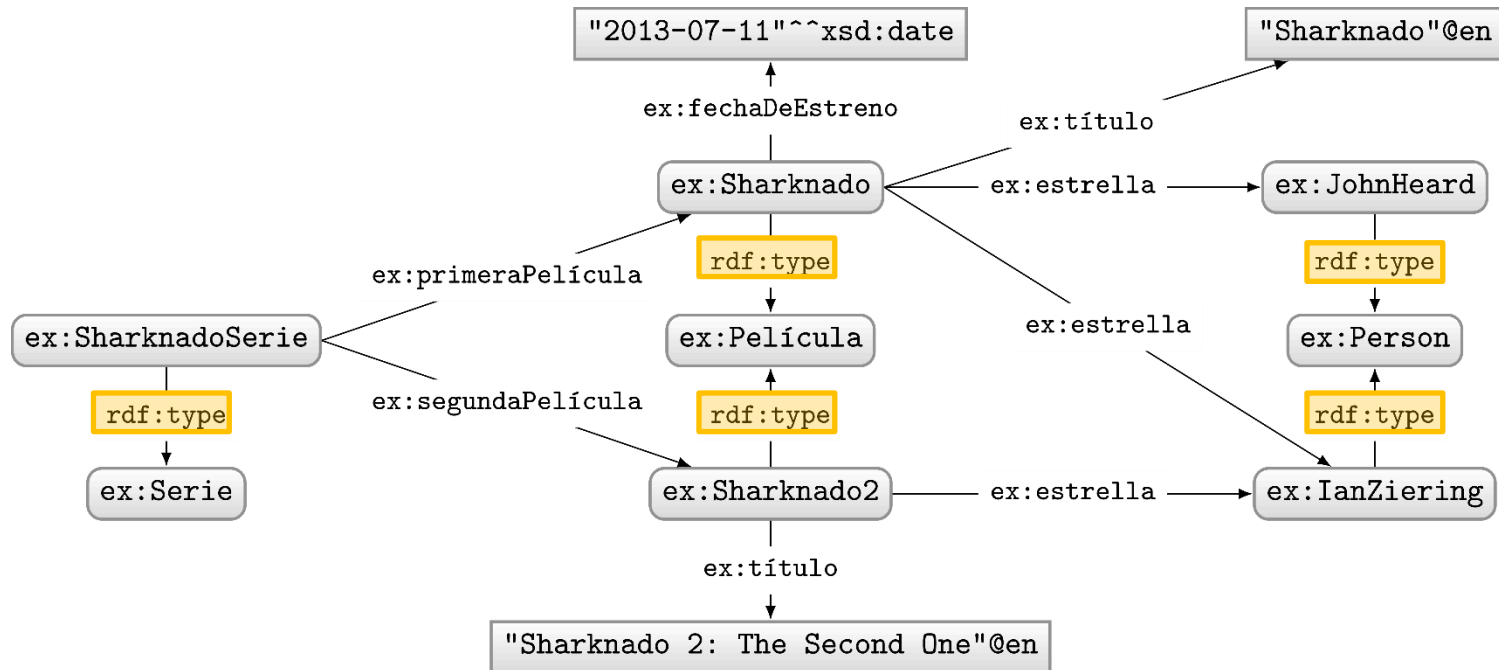


```
PREFIX ex: <http://ex.org/voc#>
SELECT *
WHERE {
  ?p a ex:Película .
  ?p ex:fechaDeEstreno ?f .
  FILTER(?f > "2013-12-31"^^xsd:date
    && ?f <= "2014-12-31"^^xsd:date)
}
```

?p ?f

Resultados vacíos

SPARQL: Filtros

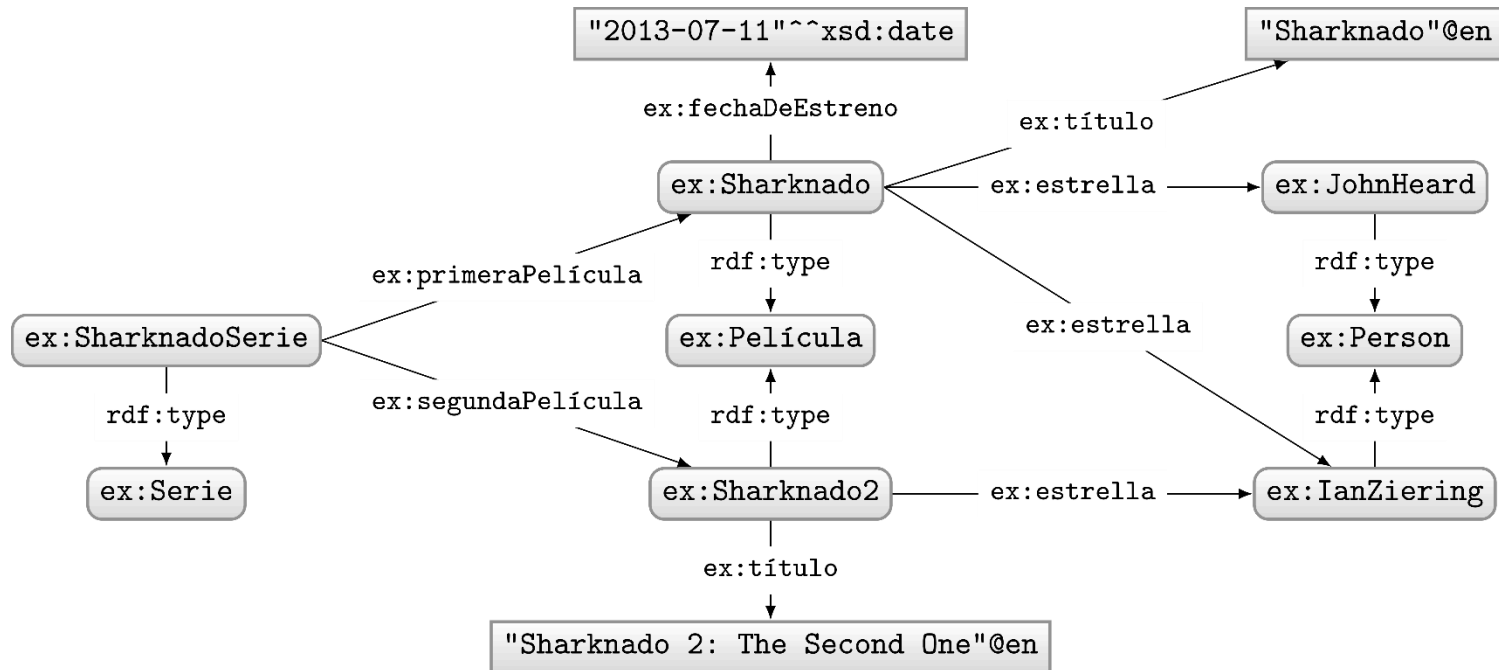


```
PREFIX ex: <http://ex.org/voc#>
SELECT *
WHERE {
  ?p a ex:Película .
  ?p ex:fechaDeEstreno ?f .
  FILTER(?f > "2013-12-31"^^xsd:date
    && ?f <= "2014-12-31"^^xsd:date)
}
```

?p	?f
----	----

Una abreviatura

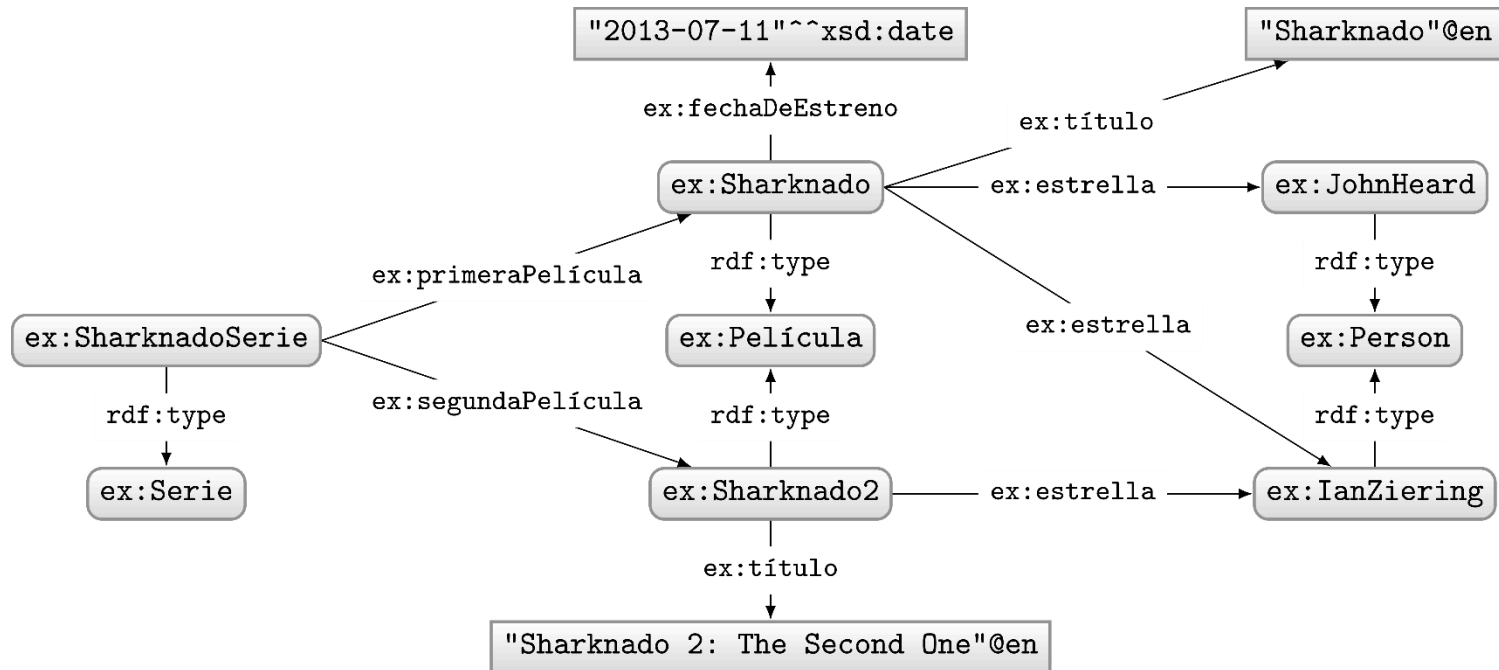
SPARQL: clausula de WHERE (otro ejemplo)



```
PREFIX ex: <http://ex.org/voc#>
SELECT *
WHERE {
  { ex:SharknadoSerie ex:primeraPelícula ?p . }
  UNION
  { ex:SharknadoSerie ex:segundaPelícula ?p . }
  OPTIONAL
  { ?p ex:fechaDeEstreno ?f . }
  ?p ex:título ?t .
  FILTER(REGEX(STR(?t), "[0-9]*"))
}
```

?p	?t	?f
ex:Sharknado2	"Sharknado 2: The Second One"@en	

SPARQL: clausula de WHERE (otro ejemplo)

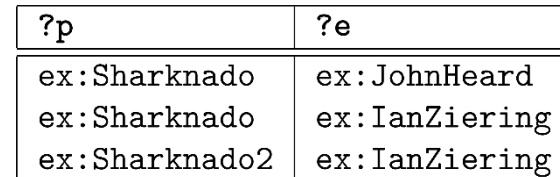


```
PREFIX ex: <http://ex.org/voc#>
SELECT *
WHERE {
    ?p a ex:Película .
    OPTIONAL
    { ?p ex:fechaDeEstreno ?f . }
    FILTER(!BOUND(?f))
}
```

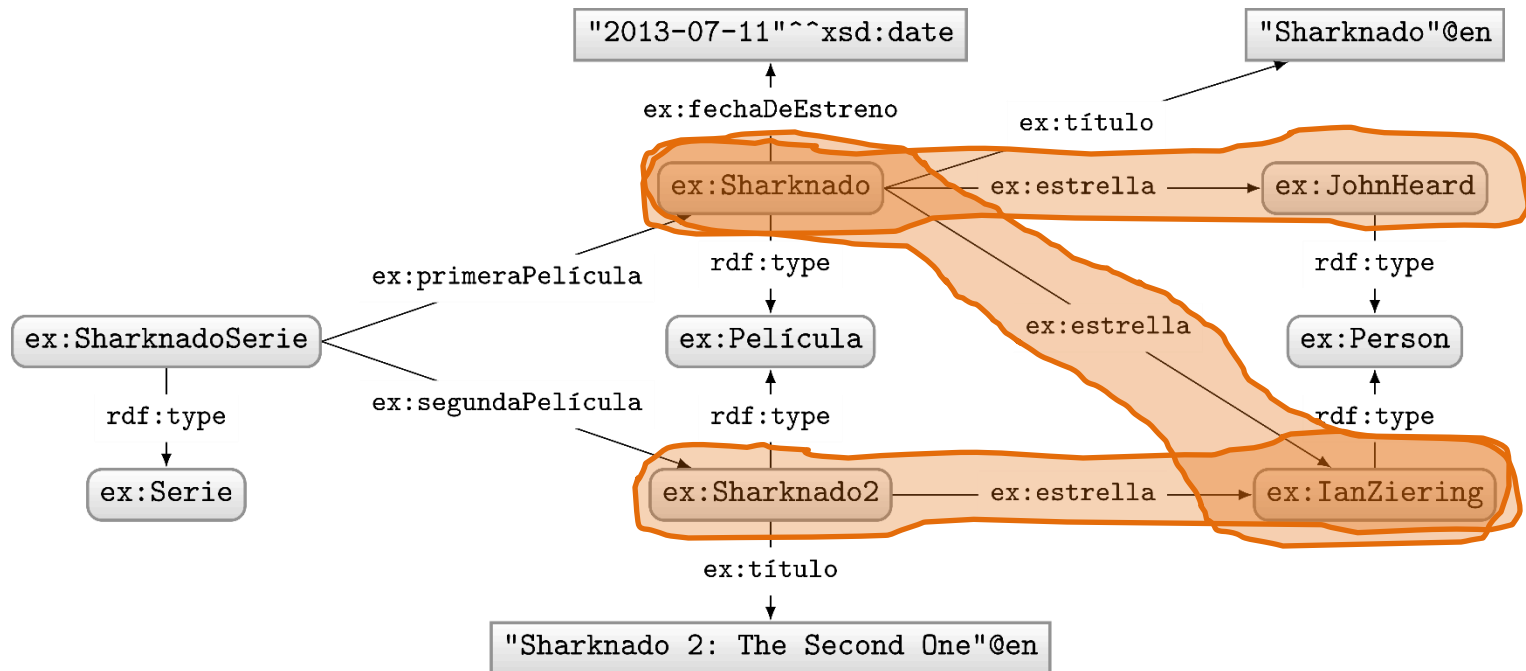
?p	?f
ex:Sharknado2	

Como NOT EXISTS/EXCEPT!

```
PREFIX ex: <http://ex.org/voc#>
SELECT *
WHERE {
    ?p a ex:Película .
    ?p ex:estrella ?e .
}
```



SPARQL: **SELECT** con proyección



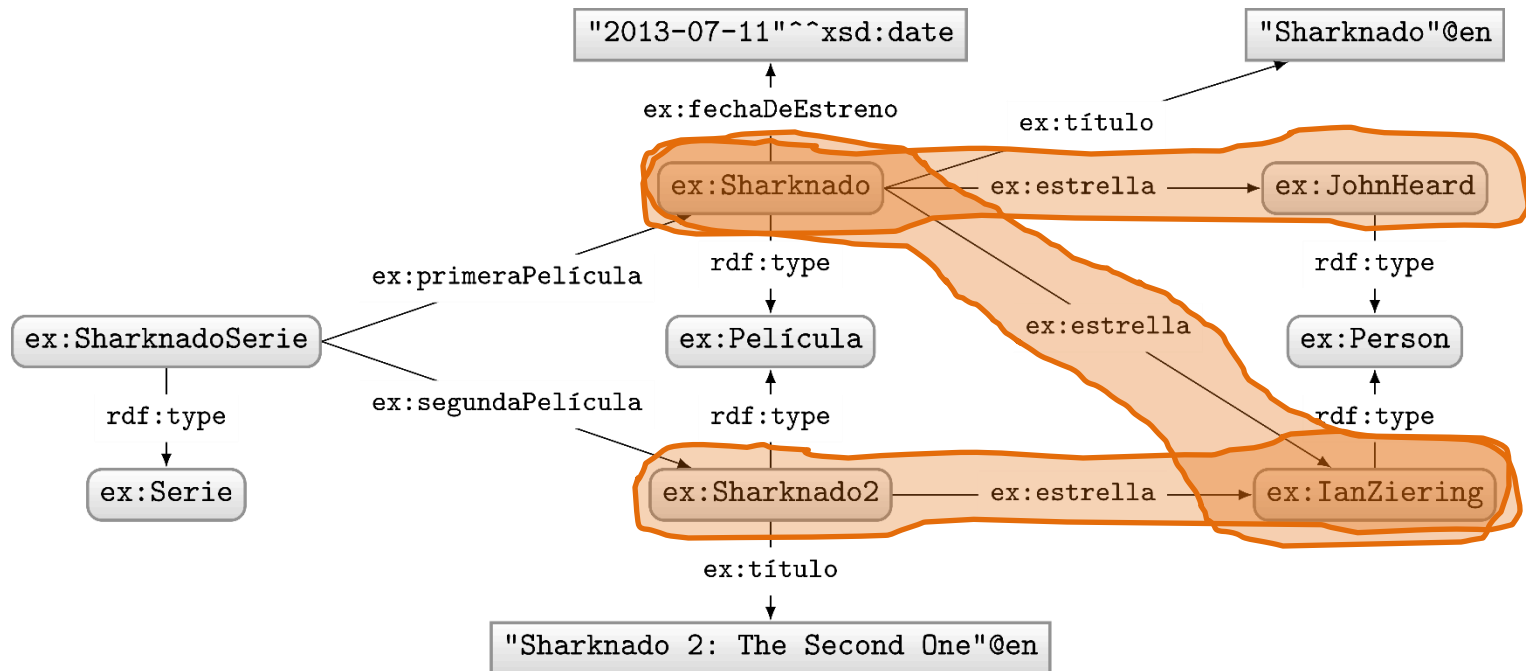
```
PREFIX ex: <http://ex.org/voc#>
SELECT ?e
WHERE {
  ?p a ex:Película .
  ?p ex:estrella ?e .
}
```

?e

ex:JohnHeard
ex:IanZiering
ex:IanZiering

Devuelve
duplicados

SPARQL: SELECT con DISTINCT

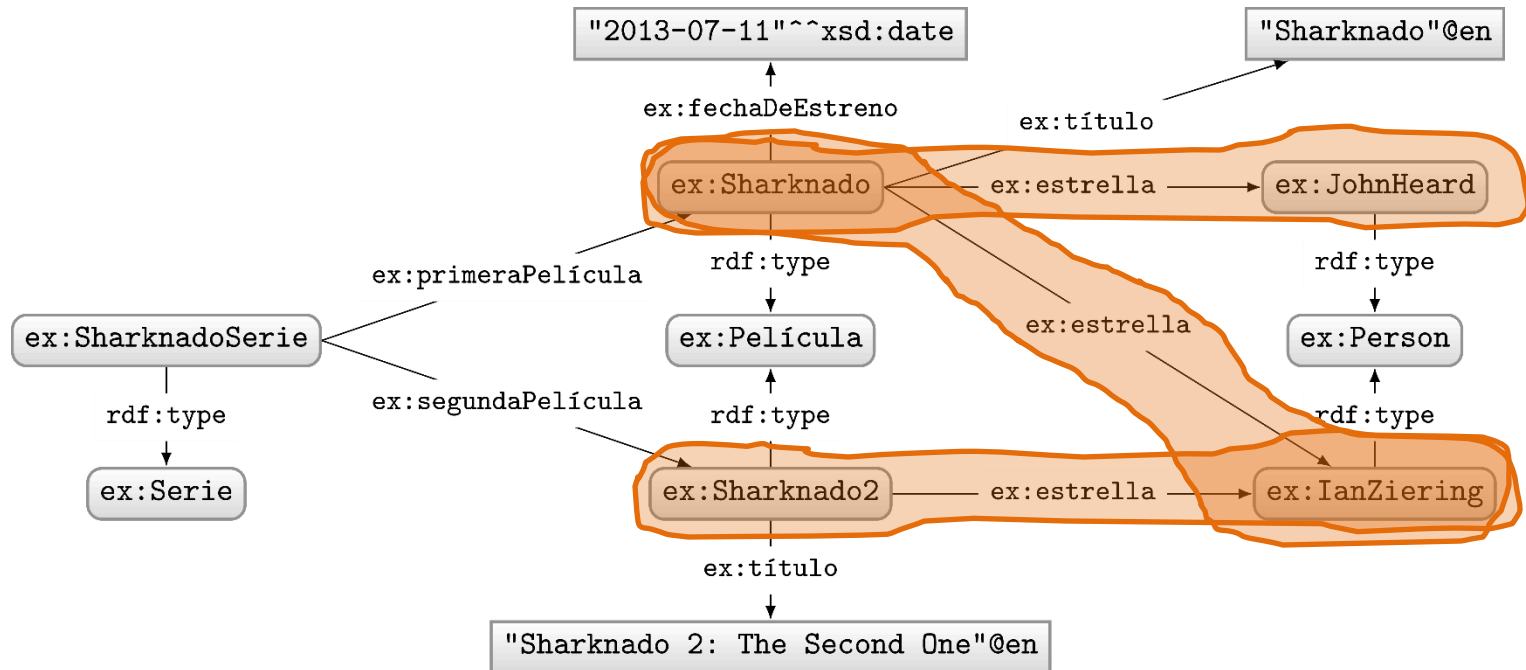


```
PREFIX ex: <http://ex.org/voc#>
SELECT DISTINCT ?e
WHERE {
  ?p a ex:Película .
  ?p ex:estrella ?e .
}
```

?e
ex:JohnHeard
ex:IanZiering

(no hay duplicados)

SPARQL: ASK

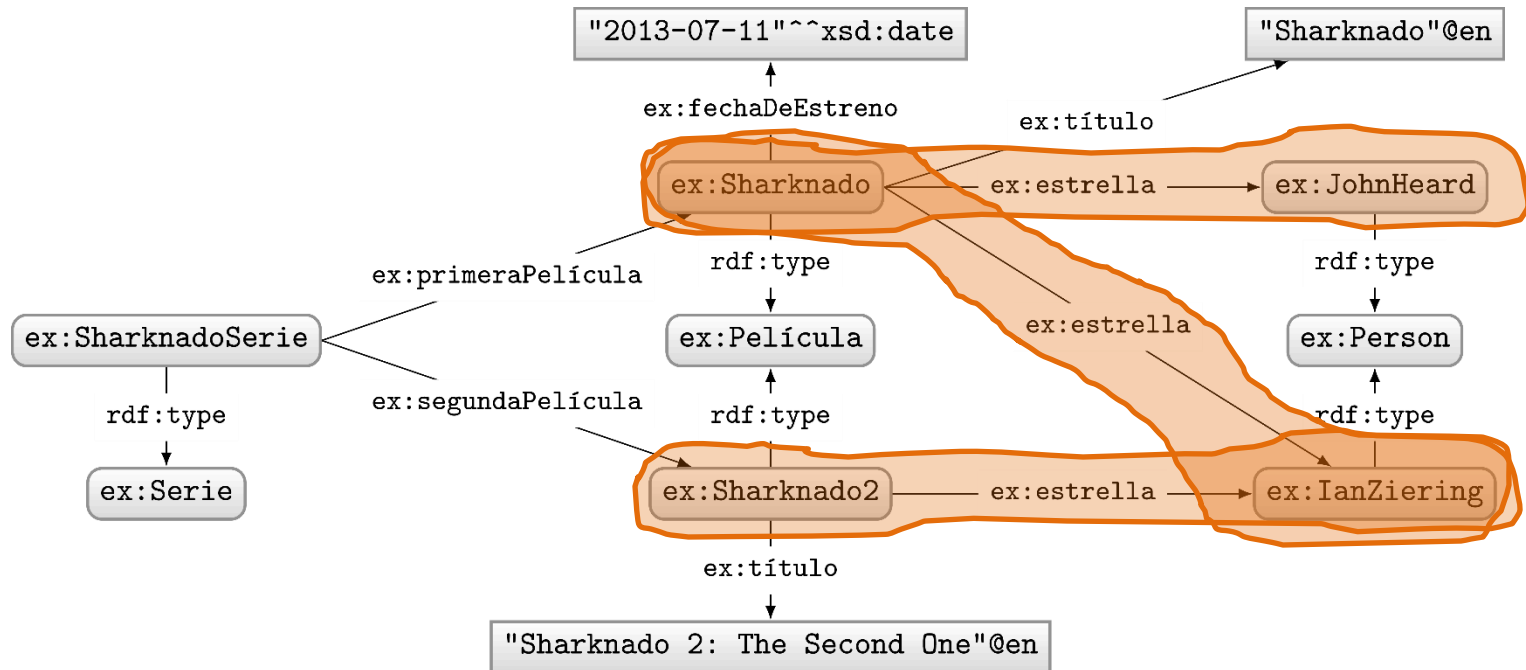


```
PREFIX ex: <http://ex.org/voc#>
ASK
WHERE {
  ?p a ex:Película .
  ?p ex:estrella ?e .
}
```

true

true si hay al menos
un resultado,
false si no.

SPARQL: CONSTRUCT



```
PREFIX ex: <http://ex.org/voc#>
CONSTRUCT { ?e ex:empleo ex:Actor }
WHERE {
  ?p a ex:Película .
  ?p ex:estrella ?e .
}
```

```
@prefix ex: <http://ex.org/voc#> .
ex:JohnHeard ex:empleo ex:Actor .
ex:IanZiering ex:empleo ex:Actor .
```

Devuelve un grafo de RDF

Modificadores: ORDER BY, LIMIT, OFFSET

Consulta: “La segunda película y la tercera película más recientes”

```
PREFIX ex: <http://ex.org/voc#>
SELECT ?p
WHERE { ?p ex:fechaDeEstreno ?f . }
ORDER BY DESC(?f)
LIMIT 2
OFFSET 1
```

Expresiones de Caminos: “*Property paths*”

e es definido recursivamente como

p	un predicado simple
$\hat{e}$	un camino inverso
e_1/e_2	e_1 seguido por e_2
$e_1 e_2$	e_1 o e_2
e^*	cero o más de e
e^+	uno o más de e
$e?$	cero o uno de e
$!p$	no p
(e)	paréntesis indican precedencia

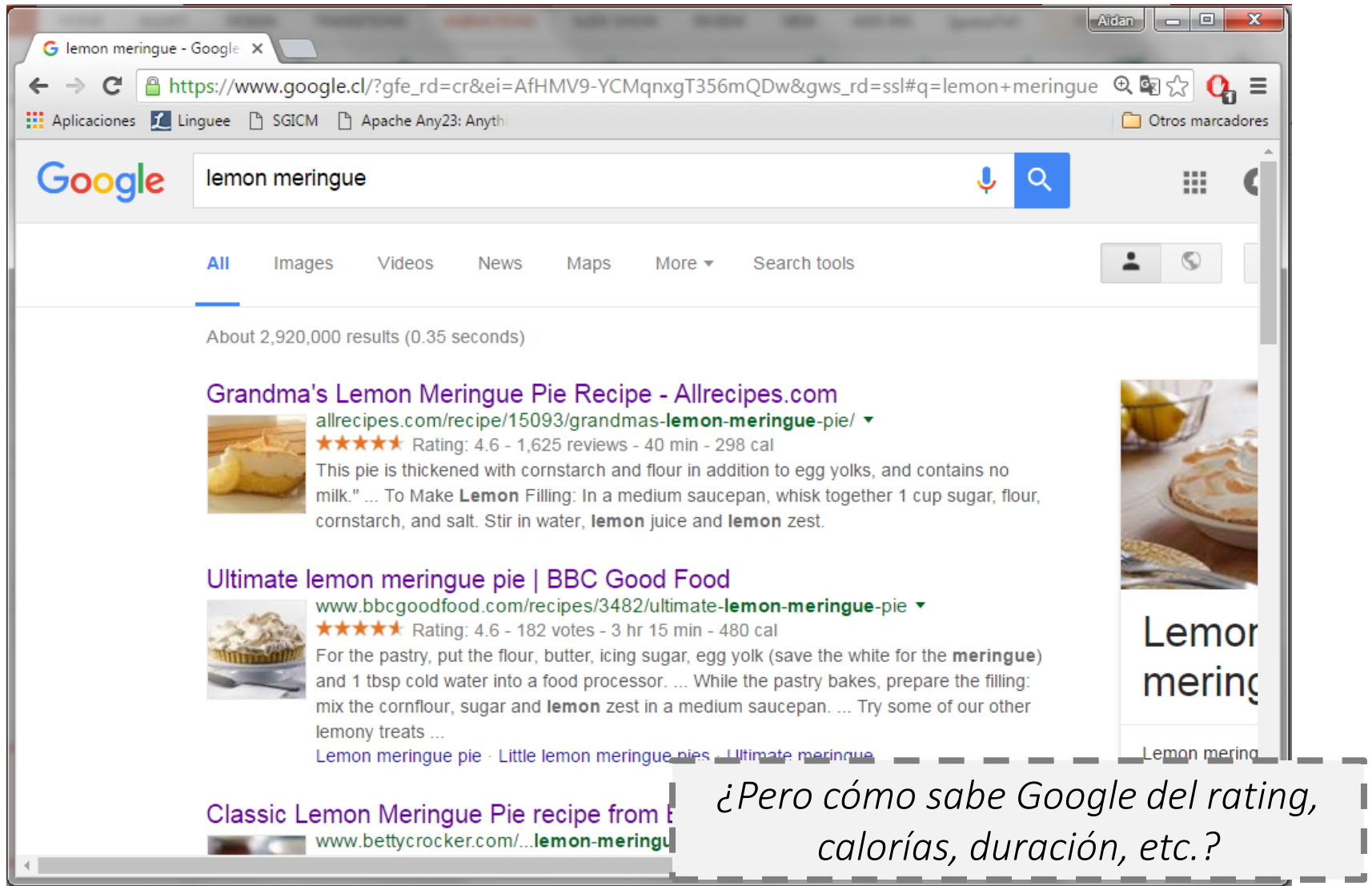
Consulta: “Los actores con un ‘Número de Bacon’ finito”

```
PREFIX ex: <http://ex.org/voc#>
SELECT ?k
WHERE {
    ex:KevinBacon (ex:estrella/^ex:estrella)* ?k .
}
```

LA WEB DE DATOS... HOY EN DÍA

GOOGLE'S "*RICH SNIPPETS*"

Resultados lujosos ...



The screenshot shows a Google search for "lemon meringue" on a desktop browser. The search results are displayed under the "All" tab. The first result is "Grandma's Lemon Meringue Pie Recipe - Allrecipes.com", which includes a small image of a pie, a star rating of 4.6, 1,625 reviews, and a duration of 40 minutes. The second result is "Ultimate lemon meringue pie | BBC Good Food", also with a 4.6 rating, 182 votes, and a duration of 3 hours 15 minutes. A third result, "Classic Lemon Meringue Pie recipe from Betty Crocker", is partially visible at the bottom. On the right side of the search results, there is a vertical strip showing a larger image of a pie and the text "Lemon meringue".

lemon meringue - Google X

https://www.google.cl/?gfe_rd=cr&ei=AfHmV9-YCMqnxgT356mQDw&gws_rd=ssl#q=lemon+meringue

Aplicaciones Linguee SGICM Apache Any23: Anyth Otros marcadores

Google lemon meringue

All Images Videos News Maps More Search tools

About 2,920,000 results (0.35 seconds)

Grandma's Lemon Meringue Pie Recipe - Allrecipes.com
allrecipes.com/recipe/15093/grandmas-lemon-meringue-pie/ ★★★★★ Rating: 4.6 - 1,625 reviews - 40 min - 298 cal
This pie is thickened with cornstarch and flour in addition to egg yolks, and contains no milk." ... To Make **Lemon** Filling: In a medium saucepan, whisk together 1 cup sugar, flour, cornstarch, and salt. Stir in water, **lemon** juice and **lemon** zest.

Ultimate lemon meringue pie | BBC Good Food
www.bbcgoodfood.com/recipes/3482/ultimate-lemon-meringue-pie/ ★★★★★ Rating: 4.6 - 182 votes - 3 hr 15 min - 480 cal
For the pastry, put the flour, butter, icing sugar, egg yolk (save the white for the **meringue**) and 1 tbsp cold water into a food processor. ... While the pastry bakes, prepare the filling: mix the cornflour, sugar and **lemon** zest in a medium saucepan. ... Try some of our other lemony treats ...
Lemon meringue pie · Little lemon meringue pies · Ultimate meringue

Classic Lemon Meringue Pie recipe from Betty Crocker
www.bettycrocker.com/...lemon-meringue

Lemon meringue

¿Pero cómo sabe Google del rating, calorías, duración, etc.?

Datos anidados en los documentos

Ultimate lemon meringue pie | BBC Good Food



www.bbcgoodfood.com/recipes/3482/ultimate-lemon-meringue-pie ▼
★★★★★ Rating: 4.6 - 182 votes - 3 hr 15 min - 480 cal

For the pastry, put the flour, butter, icing sugar, egg yolk (save the white for the meringue) and 1 tbsp cold water into a food processor. ... While the pastry bakes, prepare the filling: mix the cornflour, sugar and **lemon** zest in a medium saucepan. ... Try some of our other lemony treats ...

[Lemon meringue pie](#) - [Little lemon meringue pies](#) - [Ultimate meringue](#)

```
...  
<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>
```

- ✓ La pública recibe más clics
- ✓ Google puede crear resultados lujosos

schema.org (Bing, Google, Yahoo!, Yandex)

schema.rdfs.org

[Home](#)

[Learn](#)

[Tools](#)

[Mappings](#)

[FAQ](#)

[About](#)

What is Schema.RDFS.org?

In early June 2011, the three big search engines Bing, Google and Yahoo! introduced **Schema.org**, a collection of terms that webmasters can use to markup their pages to improve the display of search results. This site is a complementary effort by people from the Linked Data community to support Schema.org deployment and usage with a special focus on **Linked Data**:

- We provide markup **examples** and **tutorials** about publishing & consuming data with Schema.org terms.
- We maintain **mappings** from Web Data vocabularies such as the **DBpedia ontology** to Schema.org terms.
- We list **tools** and libraries that are able to consume or produce Schema.org-based data.
- We automatically **scrape** the Schema.org terms on a daily basis and generate the following formats:



RDF/Turtle



RDF/XML



RDF/NTriples



JSON



CSV: all classes, all properties

Note that the official **OWL** version of the terms is directly maintained at Schema.org and independent from the above presented formats. Various **tools** that use or produce Schema.org terms are already available or in preparation. Stay tuned!

GOOGLE'S KNOWLEDGE GRAPH

Google: "Info-box"

The screenshot shows a Google search for "sully prudhomme". The search bar contains the text "sully prudhomme". Below the search bar, the results are displayed. The first result is from Wikipedia, titled "Sully Prudhomme - Wikipedia, the free encyclopedia". The second result is from Wikipedia in Spanish, titled "Sully Prudhomme - Wikipedia, la enciclopedia libre". The third result is from Nobelprize.org, titled "Sully Prudhomme - Biographical - Nobelprize.org". The fourth result is from Nobelprize.org, titled "Sully Prudhomme - Nobelprize.org". The fifth result is from Britannica.com, titled "Sully Prudhomme | French poet | Britannica.com". The sixth result is from NNDB.com, titled "Sully Prudhomme - NNDB.com".

Google sully prudhomme

https://www.google.cl/?gfe_rd=cr&ei=AfHVMV9-YC

Aplicaciones Linguee SGICM Apache Any23: Anyth

Google sully prudhomme

All Images Videos Books News More Search tools

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 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 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. ... 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 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, died 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 Parnasse, to which Sully Prudhomme contributed several ...

Sully Prudhomme | Definition of Sully Prudhomme by Merriam-Webster

The info-box for "Sully Prudhomme" features a grid of eight portraits of the poet and other related figures. The main title "Sully Prudhomme" is prominently displayed, followed by the word "Poet". A brief biography states: "René François Armand Prudhomme was a French poet and essayist. He was the first ever winner of the Nobel Prize in Literature in 1901." Key details include his birth (March 16, 1839, Paris, France) and death (September 6, 1907, Châtenay-Malabry, France). His book "Les vaines tendresses" and his Nobel Prize in Literature are also mentioned. A section titled "People also search for" lists five other poets: Leconte de Lisle, Theodor Mommsen, Frédéric Mistral, Paul Verlaine, and Gabriel Fauré. A "More images" link is visible in the top right corner of the portrait grid.

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](#)

People also search for [View 10+ more](#)

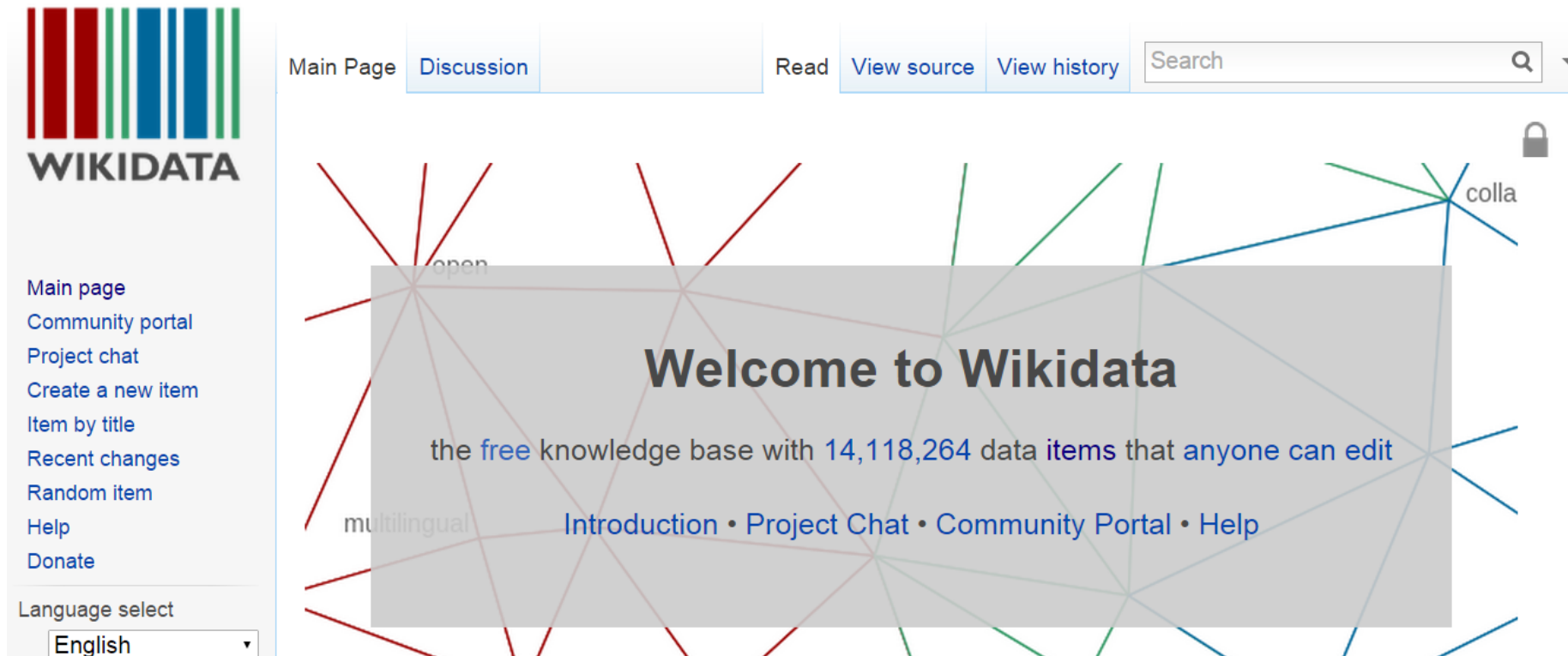
Leconte de Lisle Theodor Mommsen Frédéric Mistral Paul Verlaine Gabriel Fauré

Datos estructurados como grafos ...



LA WIKIPEDIA DE DATOS: WIKIDATA

¿Qué es Wikidata?



The image shows the Wikidata Main Page. At the top left is the Wikidata logo, consisting of a series of vertical bars in red, green, and blue, with the word "WIKIDATA" below it. To the right of the logo is a navigation bar with tabs for "Main Page", "Discussion", "Read", "View source", and "View history". A search bar is located to the right of these tabs. Below the navigation bar is a large, semi-transparent gray box with the text "Welcome to Wikidata" in bold. Below this text is the phrase "the free knowledge base with 14,118,264 data items that anyone can edit". At the bottom of the gray box is a link to "Introduction • Project Chat • Community Portal • Help". To the left of the gray box is a sidebar with links to "Main page", "Community portal", "Project chat", "Create a new item", "Item by title", "Recent changes", "Random item", "Help", and "Donate". At the bottom of the sidebar is a "Language select" dropdown menu with "English" selected. A network diagram with red, green, and blue lines is overlaid on the page, with nodes labeled "open", "multilingual", and "colla".

WIKIDATA

Main Page Discussion Read View source View history Search

Welcome to Wikidata

the free knowledge base with 14,118,264 data items that anyone can edit

Introduction • Project Chat • Community Portal • Help

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

Language select
English

Wikipedia: Varios Idiomas

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

Nationalmannschaft²

2007	Chile U-20	
2006–	Chile	76 (25)

Wikipedia: Listas, etc.



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.*

Top goalscorers [\[edit \]](#)

As of September 1, 2016

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

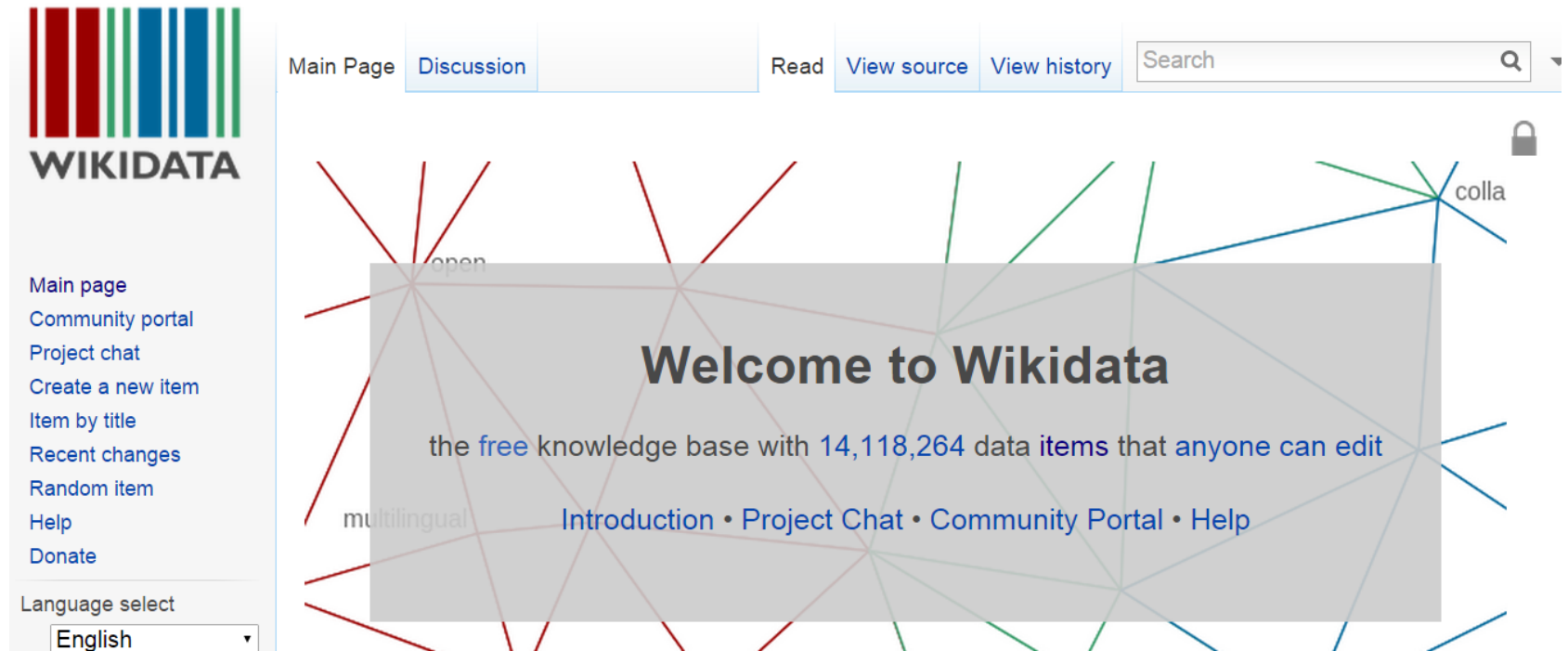
#	Name	International Career	Caps	Goals	#	Name	International Career	Goals	Caps
1.	Claudio Bravo	2004 –	106	0	1.	Marcelo Salas	1994–2007	37	70
2.	Alexis Sánchez	2006 –	102	34	2.	Iván Zamorano	1987–2001	34	69
3.	Gary Medel	2007 –	96	7		Alexis Sánchez	2006 –	34	102
4.	Gonzalo Jara	2006 –	95	3		(list)			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

Alexis marca un gol ...



*Ahora un ejército de personas tienen
que actualizar Wikipedia
(texto, listas, idiomas, etc.)*

Una solución: Wikidata



The image shows the Wikidata Main Page interface. On the left is a sidebar with the Wikidata logo and navigation links. The top navigation bar includes tabs for 'Main Page', 'Discussion', 'Read', 'View source', and 'View history', along with a search bar. The main content area features a large grey box with the text 'Welcome to Wikidata' and 'the free knowledge base with 14,118,264 data items that anyone can edit'. Below this is a link to 'Introduction • Project Chat • Community Portal • Help'. A network diagram is overlaid on the page, consisting of a central grey rectangle and numerous lines radiating outwards to various points labeled with terms like 'open', 'multilingual', 'colla', and 'colla'.

WIKIDATA

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

Language select
English

Main Page Discussion Read View source View history Search

Welcome to Wikidata
the free knowledge base with 14,118,264 data items that anyone can edit
Introduction • Project Chat • Community Portal • Help

open
multilingual
colla
colla

Wikidata: datos estructurados



Abraham Lincoln (Q91)

American politician, 16th President of the United States in office from 1861 to 1865

[Abe Lincoln](#) | [Lincoln](#) | [Honest Abe](#)

[► In more languages](#)

place of death



[Petersen House](#)

[\[edit\]](#)

[► 1 reference](#)

position held



[President of the United States of America](#)

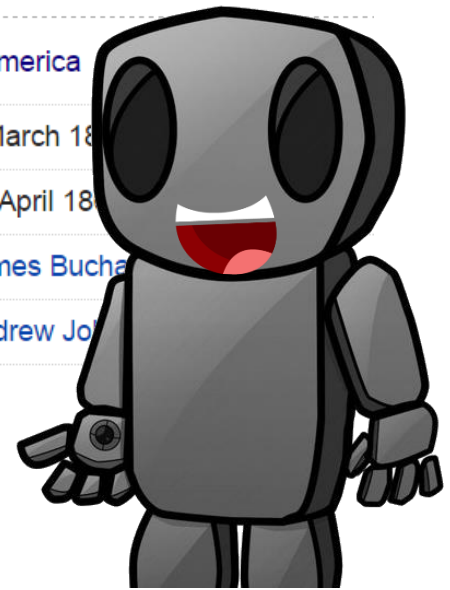
[start time](#) [4 March 1861](#)

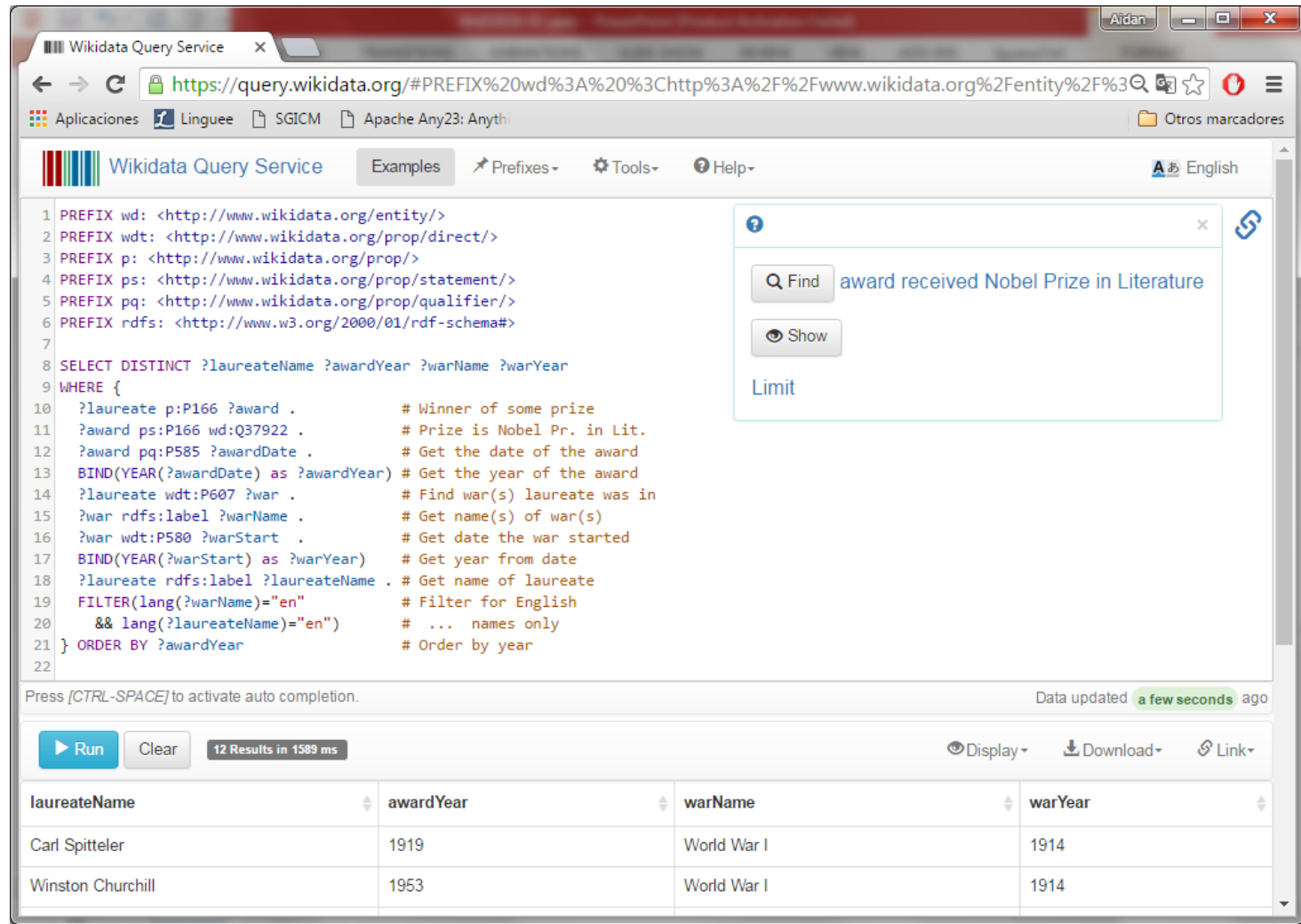
[end time](#) [15 April 1865](#)

[replaces](#) [James Buchanan](#)

[succeeded by](#) [Andrew Johnson](#)

[► 1 reference](#)





Wikidata Query Service

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

Aplicaciones

Linguee

SGICM

Apache Any23: Anythi

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

EN CONCLUSIÓN

Una área de investigación aquí ...



Datos $\neq$ Datos Relacionales



¿Preguntas?

