

Ejercicio. Consultas SPARQL en DBPEDIA

En la siguiente dirección [<http://dbpedia.org/sparql>] está el Endpoint para consultar la dbpedia. Una vez que haya accedido a la interfaz de consultas, puede proceder a realizar las siguientes consultas. Para facilitar el proceso en los recuadros encontrará los prefijos necesarios y el modelo dela consulta. Debe completar la consulta en la zona marcada como [COMPLETAR] .

1. Se quiere obtener información sobre el concepto **Gmail**, en concreto saber la homepage, el tipo de recurso asociado, los idiomas en los que está el recurso, el propietario y el creador del sitio web <https://mail.google.com>

Infobox de Gmail en Wikipedia

URL	mail.google.com
Type of site	Webmail
Available language(s)	número de idiomas en el que está
Owner	Google
Created by	Paul Buchheit

Para completar la consulta necesitarás utilizar los datos de la siguiente tabla:

type	<i>dbpedia2:type</i>
Language	<i>dbpedia2:language</i>
webpage	<i>foaf:homepage</i>
Owner	<i>dbpedia2:owner</i>
Author	<i>dbpedia2:author</i>

Consulta a completar:

PREFIX res: <<http://dbpedia.org/resource/>>

PREFIX dbpedia2: <<http://dbpedia.org/ontology/>>

PREFIX dbpedia: <<http://dbpedia.org/>>

Select ?pagina ?tipo ?idioma ?propietario ?creador where {

[COMPLETAR]

}

Para saber cómo está organizado DBpedia se necesita visitar el sitio <http://dbpedia.org/page/Gmail>



dbo:author	<ul style="list-style-type: none"> dbr:Paul_Buchheit
dbo:owner	<ul style="list-style-type: none"> dbr:Google
dbo:type	<ul style="list-style-type: none"> dbr:Webmail
dbo:wikiPageExternalLink	<ul style="list-style-type: none"> https://mail.google.com http://m.gmail.com/ https://apps.google.com/products/gmail/ https://gmail.com/
dbo:wikiPageID	<ul style="list-style-type: none"> 569005 (xsd:integer)
dbo:wikiPageRevisionID	<ul style="list-style-type: none"> 744668617 (xsd:integer)
dbp:alexa	<ul style="list-style-type: none"> Gmail.com: 4,928
dbp:tb	<ul style="list-style-type: none"> Gmail
dbp:caption	<ul style="list-style-type: none"> A screenshot of a Gmail inbox and compose box in September 2015
dbp:commercial	<ul style="list-style-type: none"> Yes

En esta ocasión se ha facilitado el nombre de los elementos de la tripleta en la tabla de más arriba. Pero con los datos de <http://dbpedia.org/page/Gmail> vemos que existe un recurso

<Gmail > cuyos otros elementos de la tripleta son

<dbo: owner><dbr:Google>, indica el propietario de Gmail es Google

<dbo:author><dbr:Paul_Buchheit>, indica que el autor de Gmail es Paul Buchheit

Dbo y dbr son prefijos estándar, las equivalencias estándar de los namespaces se pueden encontrar en el terminal endpoint <https://dbpedia.org/sparql> en la pestaña Namespace Prefixes

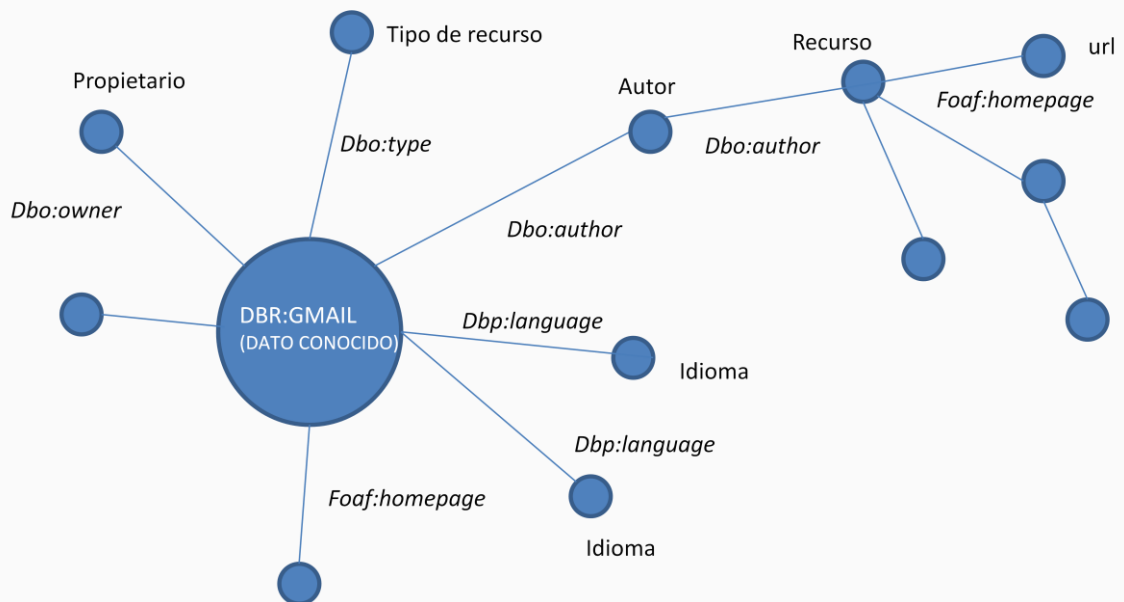
```

Select ?webpage ?type ?owner ?author where {
  dbr:Gmail foaf:homepage ?webpage.
  dbr:Gmail dbo:type ?type.
  dbr:Gmail dbo:owner ?owner.
  dbr:Gmail dbo:author ?author
}
    
```



URI	Label
http://www.w3.org/2001/sw/DataAccess/tests/test-dawg#	dawgt
http://dbpedia.org/resource/Category:	dbc
http://dbpedia.org/ontology/	dbo
http://dbpedia.org/property/	dbp
http://af.dbpedia.org/resource/	dbpedia-af
http://als.dbpedia.org/resource/	dbpedia-als
http://an.dbpedia.org/resource/	dbpedia-an
http://ar.dbpedia.org/resource/	dbpedia-ar
http://az.dbpedia.org/resource/	dbpedia-az
http://bar.dbpedia.org/resource/	dbpedia-bar
http://be.dbpedia.org/resource/	dbpedia-be
http://be-x-old.dbpedia.org/resource/	dbpedia-be-x-old

Observa que existe una dirección desde el recurso Gmail al autor o al propietario. Igual que no es lo mismo decir <el gato> <esta sobre> <el tejado>, que <el tejado> <está sobre> <el gato>. La tripleta es <dbr:gmail><dbo:autor><dbr:Paul_Buchheit> y no <dbr:Paul_Buchheit><dbo:autor><dbr:gmail>



Si dejamos los prefijos por defecto, quedaría como sigue

```
Select ?webpage ?type ?owner ?author where {
```

```
dbr:Gmail foaf:homepage ?webpage.
```

```
dbr:Gmail dbo:type ?type.
```

```
dbr:Gmail dbo:owner ?owner.
```



```
dbr:Gmail dbo:author ?author
```

```
}
```

Recuerda que los datos que ignoramos se ponen con una interrogación de cierre y un nombre aleatorio que queramos dar a ese elemento de la tripleta, por ejemplo ?author

Al ejecutar la consulta veremos lo siguiente

webpage	type	owner	author
https://mail.google.com	http://dbpedia.org/resource/Webmail	http://dbpedia.org/resource/Google	http://dbpedia.org/resource/Paul_Buchheit
https://gmail.com	http://dbpedia.org/resource/Webmail	http://dbpedia.org/resource/Google	http://dbpedia.org/resource/Paul_Buchheit

El problema de los prefijos predefinidos es que varían, por ejemplo no son los mismos en Dbpedia España que en Dbpedia inglés. Además en los últimos años han cambiado periódicamente. Una solución es poner nuestros propios prefijos de la siguiente manera.

```
PREFIX res: <http://dbpedia.org/resource/>
PREFIX dbpedia2: <http://dbpedia.org/ontology/>
Select ?webpage ?type ?owner ?author where {
res:Gmail foaf:homepage ?webpage.
res:Gmail dbpedia2:type ?type.
res:Gmail dbpedia2:owner ?owner.
res:Gmail dbpedia2:author ?author
}
```

Los nombres de los prefijos han sido puestos de forma arbitraria

Existen dos maneras de realizar la consulta, ¿se te ocurre una forma alternativa?

Recordad que conocemos dos datos:

1. El nombre del recurso se llama GMAIL
2. La ubicación del recurso es <https://mail.google.com>

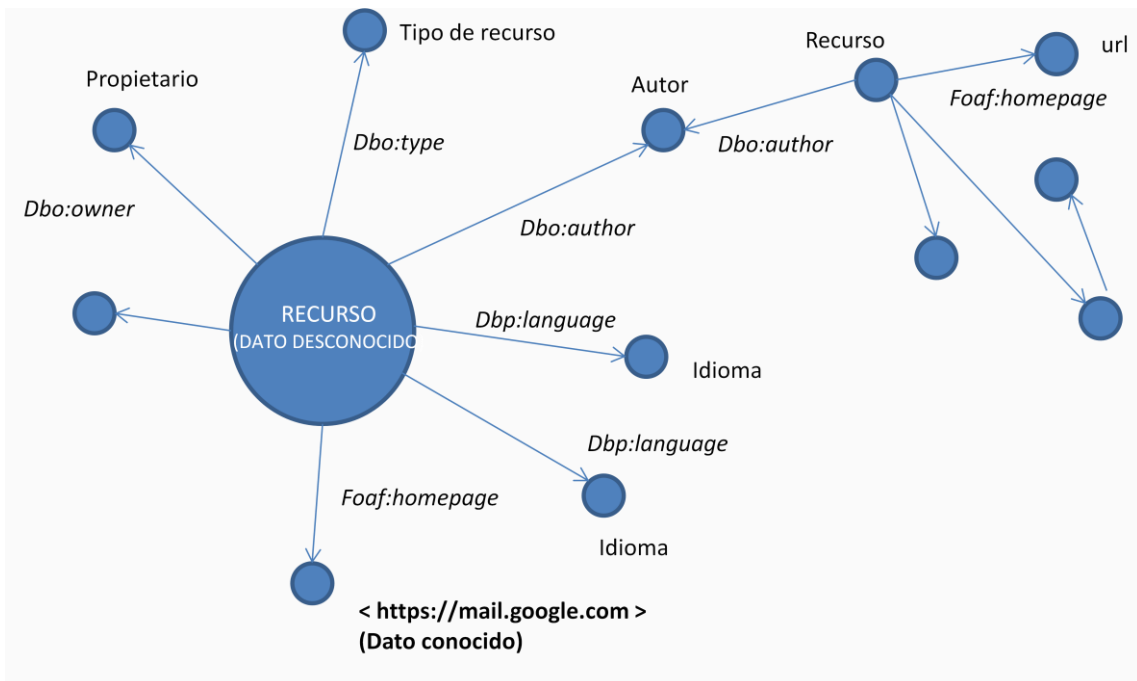
Con cualquiera de estos dos datos se puede resolver la consulta (ignorando el otro dato), esto es conocer, el tipo de recurso, el propietario y el autor.

Existe una forma alternativa de responder la consulta, si tenemos en cuenta que conocemos la página web. En esta ocasión vamos a olvidar que conocemos como se llama el recurso (Gmail) pero teniendo en cuenta, como decíamos, que conocemos la ubicación del recurso.

Para este tipo de consulta debemos tener en cuenta que en esta ocasión conocemos el último elemento de la tripleta y no el primero.

```
<recurso><foaf:homepage ><https://mail.google.com>.
```





PREFIX dbpedia2: <http://dbpedia.org/ontology/>

```
Select ?recurso ?type ?owner ?author where {
?recurso foaf:homepage <https://mail.google.com>.
?recurso dbpedia2:type ?type.
?recurso dbpedia2:owner ?owner.
?recurso dbpedia2:author ?author
}
```

El resultado es el siguiente

recurso	type	owner	author
http://dbpedia.org/resource/Gmail	http://dbpedia.org/resource/Webmail	http://dbpedia.org/resource/Google	http://dbpedia.org/resource/Paul_...

Una dificultad: Intenta incluir el número de idiomas con el que está el recurso.

Language	dbp:language
----------	--------------

Para saber el número de idiomas en el que está el recurso. Ten en cuenta que este campo habitualmente no está relleno, por lo que en ese caso no obtendrás ningún dato, ¿cómo lo pondrías para asegurarte que siguen saliendo los datos de página, tipo, propietario y creador?

Solución: en estos casos lo tendremos que poner el elemento como opcional, de no hacerlo así y no estar relleno no nos saldría ningún resultado



PREFIX dbpedia2: <<http://dbpedia.org/ontology/>>

Select ?a ?tipo ?idioma ?propietario ?creador where {{

?a dbpedia2:type ?tipo.

?a foaf:homepage <<https://mail.google.com>>.

?a dbpedia2:owner ?propietario.

?a dbpedia2:author ?creador.}

OPTIONAL

{?a dbp:language ?idioma.}}

El resultado es el siguiente:

a	tipo	idioma	propietario	creador
http://dbpedia.org/resource/Gmail	http://dbpedia.org/resource/Webmail	72	http://dbpedia.org/resource/Google	http://dbpedia.org/resource/Paul_Buchheit

2. Ahora queremos saber que recursos ha creado el el creador de Gmail (Paul Buchheit)

- Para saberla autoría el atributo es onto:author
- La URI dePaul Buchheit está en http://dbpedia.org/resource/Paul_Buchheit

Ésta es la consulta a completar:

PREFIX onto: <<http://dbpedia.org/ontology/>>

Select ?autoriaBuchheit where {

[COMPLETAR]

PREFIX onto: <<http://dbpedia.org/ontology/>>

Select ?autoriaBuchheit where {

?autoriaBuchheit onto:author <http://dbpedia.org/resource/Paul_Buchheit>.}

Estos son los resultados:

autoriaBuchheit

<http://dbpedia.org/resource/Gmail>

Quizás más que el criterio autor en este caso es relevante el criterio de la ontología knownFor <http://live.dbpedia.org/ontology/knownFor>

PREFIX onto: <<http://dbpedia.org/ontology/>>

Select ?autoriaBuchheit where {

<http://dbpedia.org/resource/Paul_Buchheit> onto:knownFor ?autoriaBuchheit.}

autoriaBuchheit
http://dbpedia.org/resource/Google
http://dbpedia.org/resource/FriendFeed



<http://dbpedia.org/resource/Gmail>

El motivo por el que salen diferentes soluciones lo podemos ver al visitar los sitios web que definen cada concepto

<http://live.dbpedia.org/ontology/knownFor>: Work, historic event, etc that the subject is known for.
Applies to Person, Organization, ConcentrationCamp, etc

al visitar <http://live.dbpedia.org/ontology/author> no hay una definición pero si unos atributos claros que nos indica que es una persona (rdfs:range) autora de un trabajo (dbo:work y rdfs:suPropertyOf dul:coparticipatesWith), además indica que existe un sinónimo en Wikidata (<https://www.wikidata.org/wiki/Property:P50>) y en este sitio se define autor como "main creator(s) of a written work (use on works, not humans)"

Es decir realmente KnownFor es que ha estado relacionado con el recurso de alguna manera y autor indica que es la autoría del recurso es suya. El que se ponga de una manera u otra a menudo depende del codificador. Hay que tener en mente que Dbpedia es una obra colaborativa producto de miles de personas.

