

Publicación y consumo de datos de un Knowledge Graph

Janneth Chicaiza
Interoperabilidad de datos
jachicaiza@utpl.edu.ec

Publicación de datos

Publicación de datos enlazados



Dataset
Publication

Vocabulary
Publication

Metadata
definition and
Publication

1. Publicación de **datos RDF**:
2. Publicación del **modelo ontológico o vocabulario** (y su documentación): Utilizar catálogos como: LOV vocabularies, GIT o repositorios especializados.
3. Publicación de **metadatos** utilizando vocabularios conocidos: utilizar vocabularios DCAT, VoID

Publicación de datos RDF

1. Almacenar los datos en un RDF triple store
2. Proveer acceso a los usuarios a través de un **SPARQL EndPoint**.
3. Utilizar un LOD client o librería para convertir RDF en HTML: [ELDA](#), [Pubby](#), [RDFlib.js](#)

SPARQL Query Editor

Extensions: cxml

Default Data Set Name (Graph IRI)
http://dbpedia.org

Query Text

```
select distinct ?Concept where {} a ?Concept} LIMIT 100
```

Results Format
HTML

Execute Query Reset

YASGUI e_WPB_ENG_20100908_00 tackbp.org/page/kb/e... RDF Kelvin

About e_WPB_ENG_20100908_0051_11

Property	Value
kbp:a_location_of	<ul style="list-style-type: none">kb:e_WPB_ENG_20100607_0046_1
kbp:births_in_country	<ul style="list-style-type: none">kb:e_WPB_ENG_20100906_0022_47kb:e_WPB_ENG_20101211_0016_11
kbp:canonical_mention	<ul style="list-style-type: none">Baltimore CityBaltimore County
kbp:canonical_name	<ul style="list-style-type: none">baltimore county
kbp:descriptions	<ul style="list-style-type: none">the country^1
kbp:first_name_mention	<ul style="list-style-type: none">BaltimoreBaltimore CityBaltimore County

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

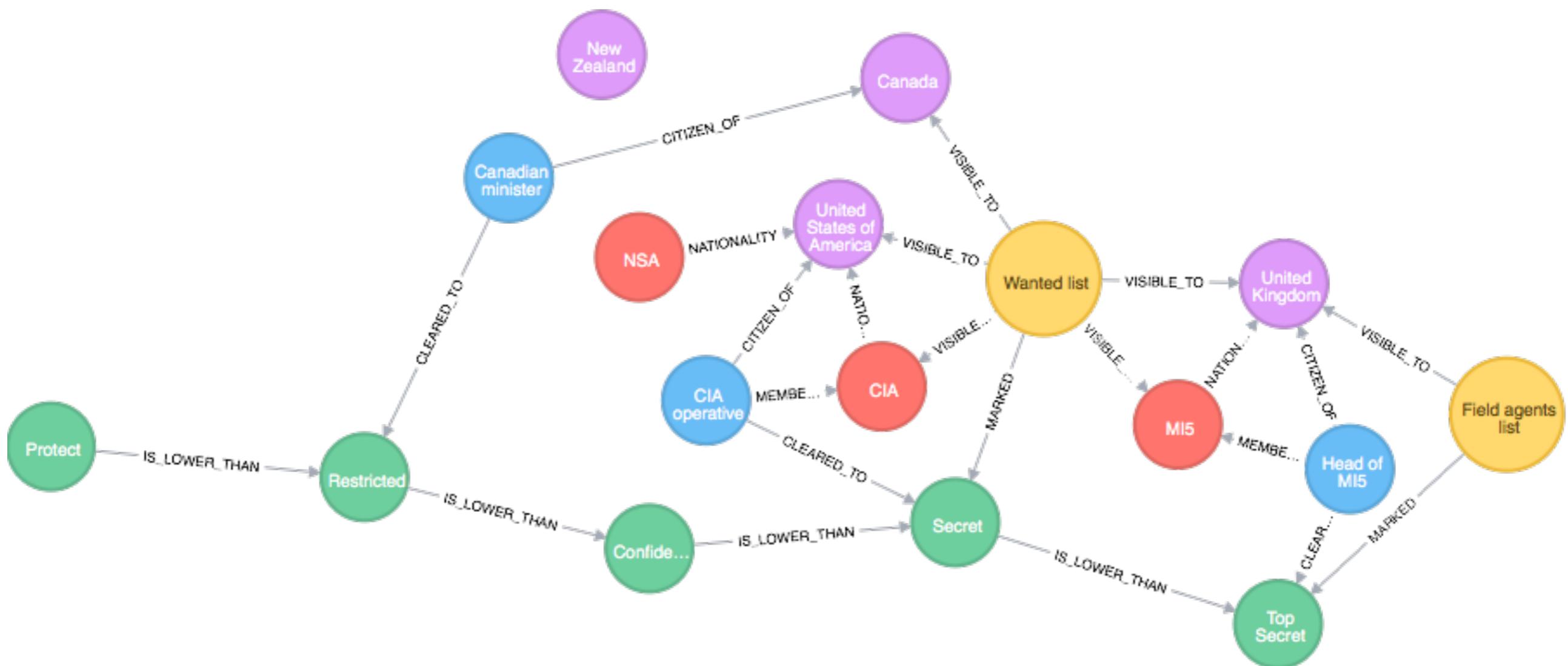
<https://github.com/cygri/pubby>

SPARQL

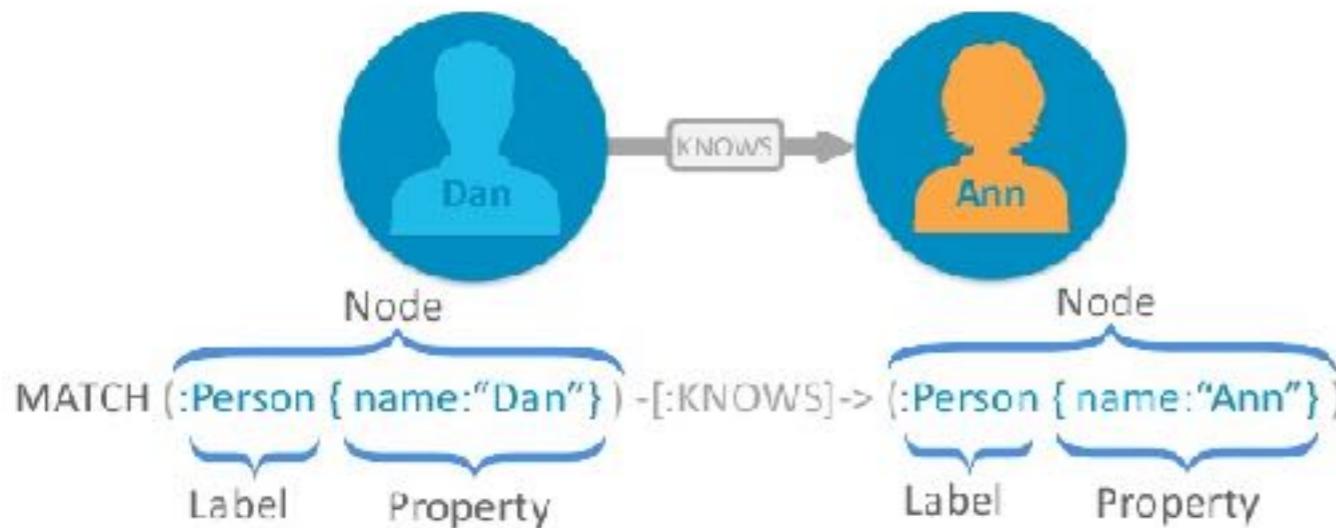


Consulta de datos

Lenguajes de consulta para datos RDF y grafos de conocimiento



Cypher Query Language



Cypher is a declarative graph query language that allows for expressive and efficient data querying in a property graph



Gremlin
 $G = (V, E)$

Gremlin is a graph traversal language. Neo4j, Giraph
Ex: g.V().label().groupCount()





Data Sources



SPARQL Protocol

Access

SPARQL
Endpoint

<https://www.w3.org/wiki/SparqlEndpoints>

SPARQL
Client/GUI

<https://yasgui.triply.cc>
<http://www.ldodds.com/projects/twinkle/>

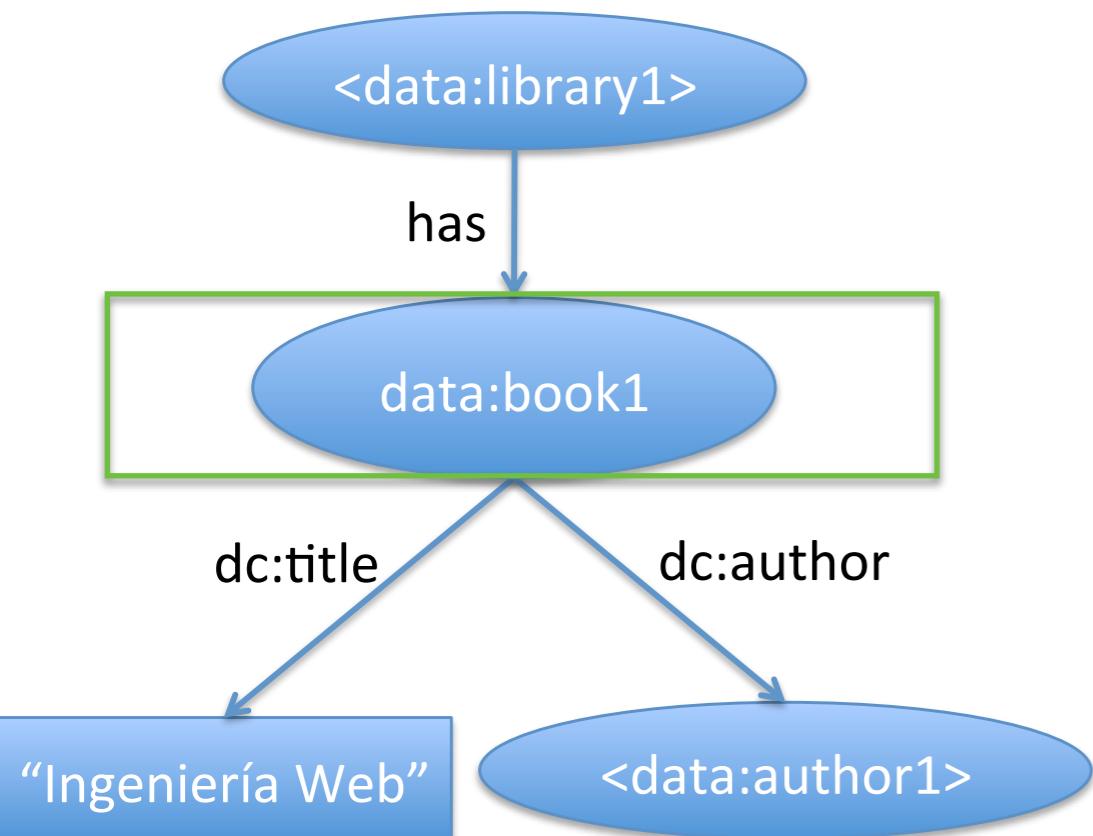
SPARQL
Libraries/
APIS

Python: SparqlWrapper

- Protocolo y lenguaje de consulta recomendado por el W3C. Permite recuperar datos RDF a través de la Web.
- Soporta cuatro formas de consulta, las cuales se basan en la coincidencia de patrones para formar conjuntos de resultados o grafos RDF: **SELECT**, **CONSTRUCT**, **ASK**, y **DESCRIBE**.
- La versión 1.1 soporta operaciones de inserción, actualización y borrado de tripletas.

SPARQL: Consulta Básica

¿Cuál es el título de data:book1?

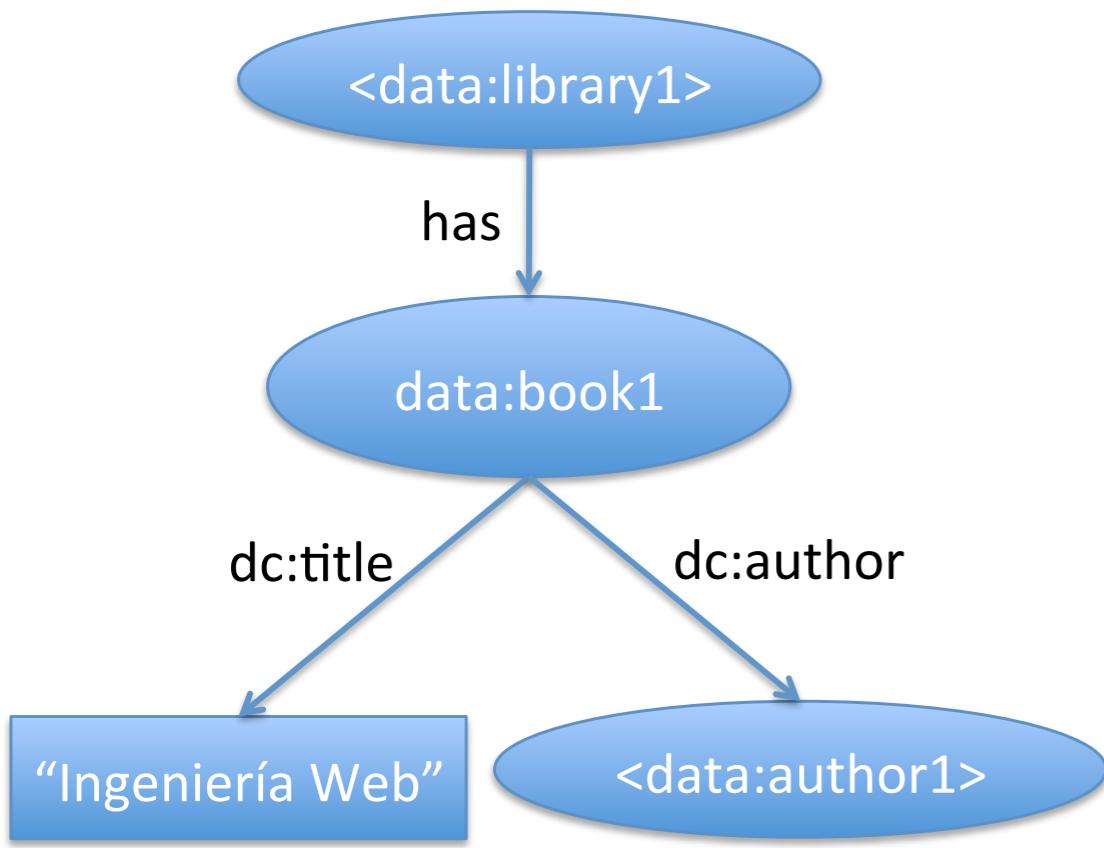


```
PREFIX dc: <http://purl.org/dc/elements/1.1/>
PREFIX data: <http://example.org/data/>
SELECT *
FROM <http://example.org/>
WHERE
{
    data:book1 dc:title ?titulo
}
```

Resultado

?titulo
"Ingeniería Web"

Diferentes patrones



data:book1 dc:title ?titulo

?titulo
“Ingeniería Web”

data:book1	?metadata	?valor
-------------------	------------------	---------------

?metadata	?valor
dc:title	“Ingeniería Web”
dc:author	data:author1

?libro	dc:title	?titulo
---------------	-----------------	----------------

?libro	?titulo
data:book1	“Ingeniería Web”

?s	?p	?o
-----------	-----------	-----------

?s	?p	?o
data:library1	:has	data:book1
data:book1	dc:title	“Ingeniería Web”
data:book1	dc:author	data:author1

Sujeto	Predicado	Objeto
<http://example.org/book/book1>	rdf:type	<http://example.org/onto/Book>
<http://example.org/book/book1>	dc:title	“Ingeniería Web”
<http://example.org/book/book1>	dc:author	<http://example.org/person/author1>
<http://example.org/person/author1>	foaf:name	“Luis Castro”
<http://example.org/person/author1>	rdf:type	<http://example.org/onto/Author>
<http://example.org/book/book2>	dc:title	“Sistemas Basados en Conocimiento”

¿Cuáles son los libros que existen en el grafo?

PREFIX dc: <http://purl.org/dc/elements/1.1/>

PREFIX onto: <http://example.org/onto/>

SELECT DISTINCT *

FROM <http://example.org/>

WHERE

{

?libro rdf:type onto:Book .

}

Resultado

?libro

<http://example.org/book/book1>

Sujeto	Predicado	Objeto
<http://example.org/book/book1>	rdf:type	<http://example.org/onto/Book>
<http://example.org/book/book1>	dc:title	“Ingeniería Web”
<http://example.org/book/book1>	dc:author	<http://example.org/person/author1>
<http://example.org/person/author1>	foaf:name	“Luis Castro”
<http://example.org/person/author1>	rdf:type	<http://example.org/onto/Author>
<http://example.org/book/book2>	dc:title	“Sistemas Basados en Conocimiento”
<http://example.org/book/book2>	rdf:type	<http://example.org/onto/Book>
<http://example.org/book/book2>	dc:author	<http://example.org/person/author2>

¿Cuál es el título de cada libro y el nombre de cada autor?

PREFIX dc: <http://purl.org/dc/elements/1.1/>

SELECT *

FROM <http://example.org/>

WHERE

{

?libro rdf:type <http://example.org/onto/Book> .

?libro dc:title ?titulo .

?libro dc:author ?autor .

OPTIONAL{

?autor foaf:name ?nombreAutor .

}

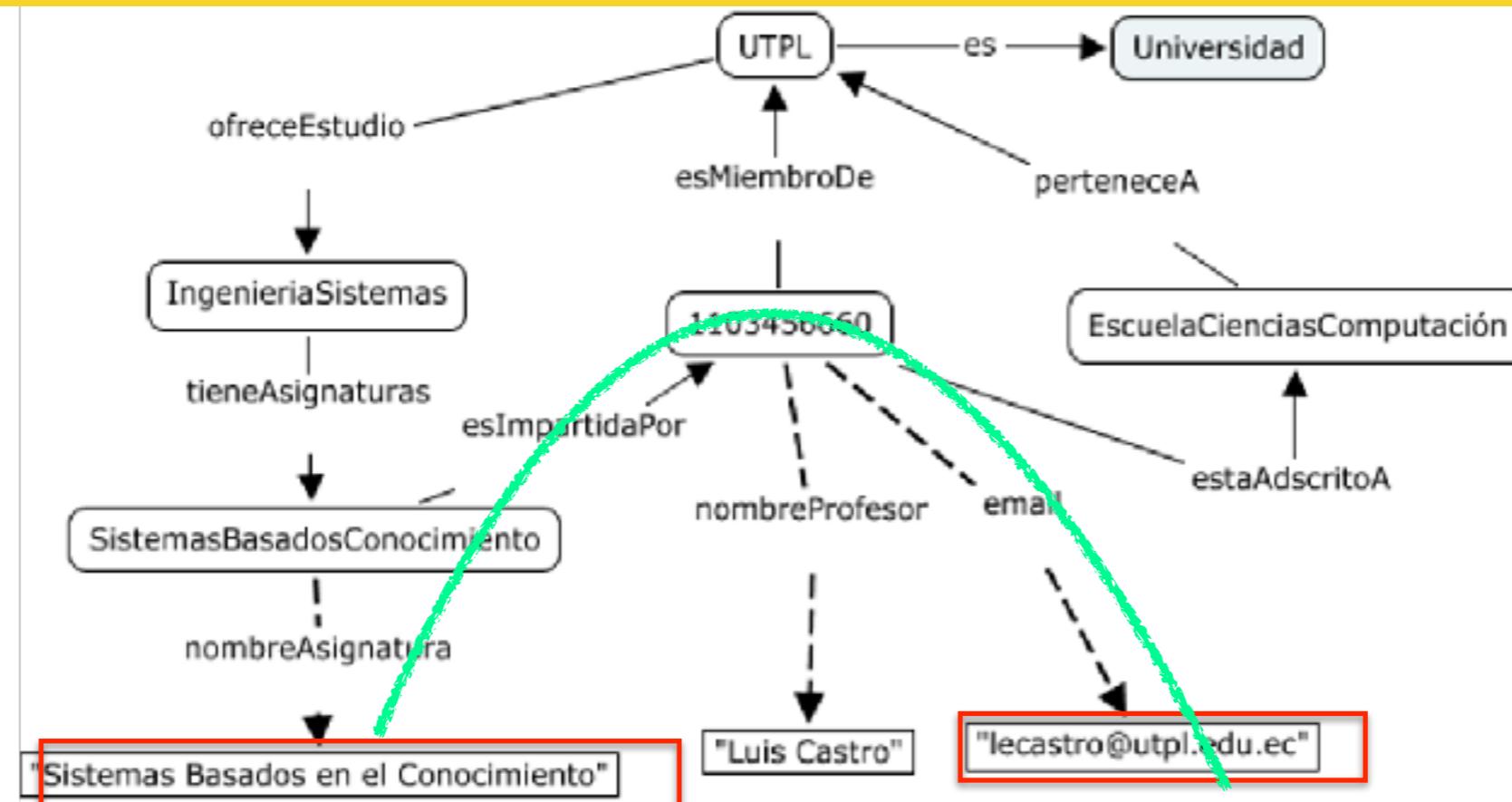
}

Resultado

?titulo	?nombreAutor
“Ingeniería Web”	“Luís Castro”
“Sistemas Basados en Conocimiento”	

¿Cuál es el nombre de los profesores que imparten asignaturas de la UTPL?

¿Cuál es el nombre de los profesores que trabajan en la UTPL?

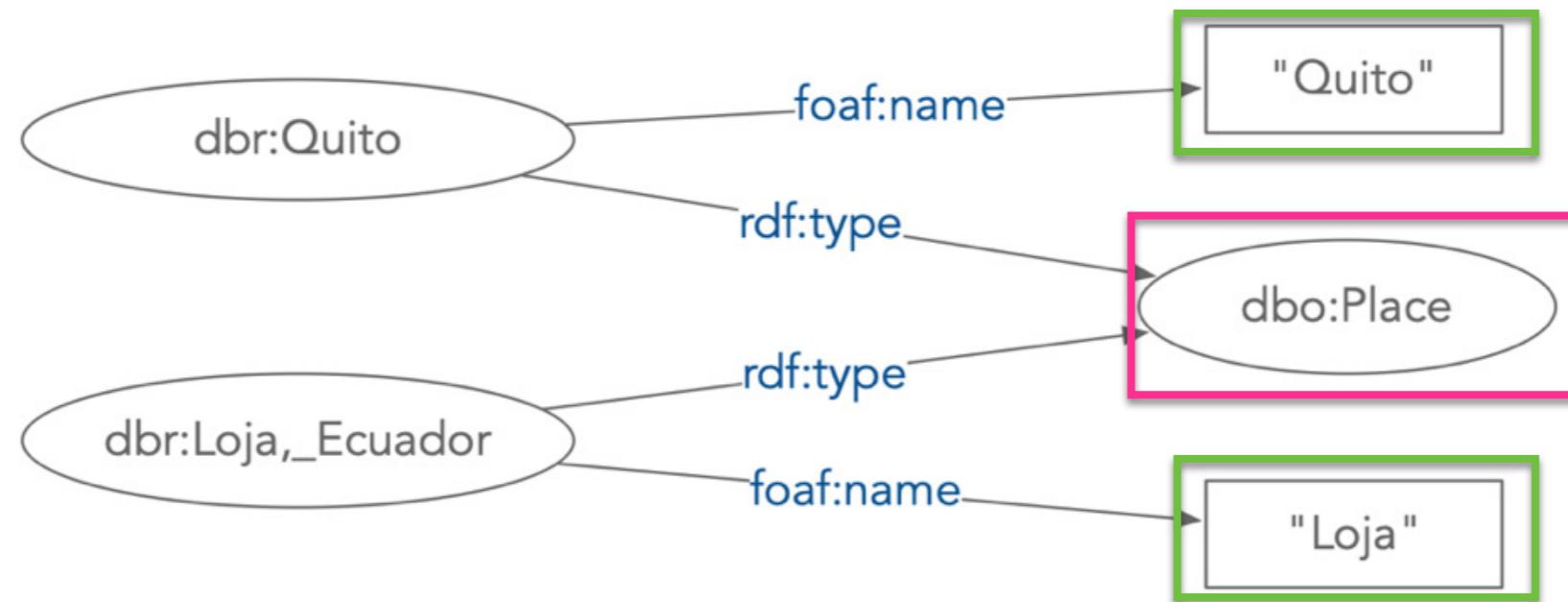


¿Cuál es el **correo electrónico** del profesor que imparte “**Sistemas Basados en el Conocimiento**”?

Ejercicios

Consulta a partir de un grafo

Considere el grafo que se indica en la figura y a partir de la estructura definida resuelva la pregunta de competencia que se indica a continuación.



CQ. ¿Cuáles son los **nombres** de las **ciudades** que aparecen en el grafo?

nombres -> foaf:name

ciudad -> dbo:Place

Consulta desde GraphDB - Grafo de Universidades



<http://example.org/universities/Ecuador>

PREFIX : <http://example.org/universities/>
 PREFIX monto: <http://example.org/ontology/>
 SELECT * WHERE {
 VALUES ?u { :UniversidaddeCuenca }
 ?u ?p ?o .
 } LIMIT 100

u	p	o
http://example.org/universities/UniversidaddeCuenca	rdf:type	foaf:Organization
http://example.org/universities/UniversidaddeCuenca	foaf:name	"Universidad de Cuenca"
http://example.org/universities/UniversidaddeCuenca	dbo:country	http://example.org/universities/Ecuador
http://example.org/universities/UniversidaddeCuenca	monto:hasDataRank	_:genid-62854dbc56744fdd972fcd5d08464afa-node1fmb2v6dhx566
http://example.org/universities/UniversidaddeCuenca	monto:hasDataRank	_:genid-62854dbc56744fdd972fcd5d08464afa-node1fmb2v6dhx567
http://example.org/universities/UniversidaddeCuenca	monto:website	http://www.ucuenca.edu.ec/

Nodos sin
URI
específica

PREFIX monto: <http://example.org/ontology/>

SELECT * WHERE {

VALUES ?u {<http://example.org/universities/UniversidaddeCuenca>}
?u monto:hasDataRank ?r .
?r ?prop ?value }

u	r	prop	value
http://example.org/universities/UniversidaddeCuenca	_:genid-62854dbc56744fdd972fcd5-d08464afa-node1fmb2v6dhw566	monto:position	"326.0"
http://example.org/universities/UniversidaddeCuenca	_:genid-62854dbc56744fdd972fcd5-d08464afa-node1fmb2v6dhw566	monto:region	"Latinoamerica"
http://example.org/universities/UniversidaddeCuenca	_:genid-62854dbc56744fdd972fcd5-d08464afa-node1fmb2v6dhw566	monto:rankingDate	"2020"
http://example.org/universities/UniversidaddeCuenca	_:genid-62854dbc56744fdd972fcd5-d08464afa-node1fmb2v6dhw566	monto:provenance	"Scimago"
http://example.org/universities/UniversidaddeCuenca	_:genid-62854dbc56744fdd972fcd5-d08464afa-node1fmb2v6dhw567	monto:position	"259"
http://example.org/universities/UniversidaddeCuenca	_:genid-62854dbc56744fdd972fcd5-d08464afa-node1fmb2v6dhw567	monto:region	"Latinoamerica"
http://example.org/universities/UniversidaddeCuenca	_:genid-62854dbc56744fdd972fcd5-d08464afa-node1fmb2v6dhw567	monto:rankingDate	"2020"
http://example.org/universities/UniversidaddeCuenca	_:genid-62854dbc56744fdd972fcd5-d08464afa-node1fmb2v6dhw567	monto:provenance	"Webometrics"

The screenshot shows the GraphDB interface with a red box highlighting the 'SPARQL' section. The main area displays a SPARQL query results table.

GraphDB Interface:

- Import:** A red icon with a downward arrow.
- Explore:** A red icon with three nodes.
- Graphs overview**
- Class hierarchy**
- Class relationships**
- Visual graph**
- Similarity**
- SPARQL** (highlighted with a red box)

SPARQL Query Results:

Showing results from 1 to 100

	nombre
1	"Universidad Católica de Santa Fe"
2	"Universidad Católica de Cuyo"
3	"Universidad de Flores"
4	"Universidad Nacional de San Luis"
5	"Universidad de la Cuenca del Plata"
6	"Universidad Nacional de La Plata"
7	"Universidad Nacional del Nordeste"
8	"Universidad Nacional del Chaco Austral"

SPARQL Query:

```
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
select distinct ?nombre where
{
    ?u a foaf:Organization;
        foaf:name ?nombre.
} limit 100
```

¿Cuál es el rank y el país de cada universidad según Webometrics?

PREFIX myonto: <<http://dbpedia.org/ontology2/>>

PREFIX monto: <<http://example.org/ontology/>>

PREFIX foaf: <<http://xmlns.com/foaf/0.1/>>

PREFIX rdf: <<http://www.w3.org/1999/02/22-rdf-syntax-ns#>>

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

select distinct ?name ?c ?pos where {

?u a foaf:Organization; foaf:name ?name ; dbo:country ?c.

?u monto:hasDataRank ?rank .

?rank monto:position ?pos; monto:provenance "Webometrics".

} limit 10

name	c	pos
"Universidad Católica de Santa Fe"	http://example.org/universities/Argentir	"1099"
"Universidad Católica de Cuyo"	http://example.org/universities/Argentir	"1107"
"Universidad de Flores"	http://example.org/universities/Argentir	"1144"
"Universidad Nacional de San Luis"	http://example.org/universities/Argentir	"119"
"Universidad de la Cuenca del Plata"	http://example.org/universities/Argentir	"1258"

Enlaces de referencia

<http://www.w3.org/TR/sparql11-query/>

<http://skos.um.es/TR/rdf-sparql-query/>



Gracias por su
Atención.