



UNIVERSIDAD TÉCNICA PARTICULAR DE LOJA

Sistemas Basados en Conocimientos

Tema: Análisis y recolección de datos

Estudiante: Alexis Fabian Montoya Parra

Juan Andrés Ramon Zhigui

Luis Miguel Rojas Paccha

Transformación de datos RDF

En esta sección se describe los pasos necesarios desde la extracción de datos sobre publicaciones COVID-19 hasta la generación de datos RDF con Jena, un framework para la construcción de ontologías.

Extracción de datos

La fuente de datos en donde se extrajo las publicaciones científicas sobre Covid-19 fue en la plataforma Crossref y Scholar Academic, una base de datos científica en donde se puede encontrar publicaciones en diferentes categorías.

Limpieza de datos

Con los datos disponibles, el siguiente paso es refinarlo para evitar problemas de caracteres con las generaciones de las tripletas RDF que puedan provocar. Para la limpieza de datos se realizó las siguientes actividades:

- Transformar el formato del archivo CSV a un formato legítimo para que pueda ser compatible con el formato de la versión de Windows 10.
- Agrupar campos en una sola celda
- Eliminar caracteres especiales de cada registro

El resultado final es un archivo mucho más preparado y estructurado listo para realizar la extracción de los datos pulidos para la creación de las tripletas el resultado se puede apreciar en la siguiente imagen.

	B	C	D	E	F	G
1	date-time	publisher	DOI	type	source	title
3	2020-10-23T17:13:59Z	The International Centre for Trade Unio	10.14213/inteuniorigh.27.3.0020	journal-article	Crossref	'Just because you don't see your boss, doesn't mean you don't have a boss': Covic
4	2020-08-14T01:40:41Z	American Association for the Advancem	10.1126/science.abe2995	journal-article	Crossref	'We're losing an entire generation of scientists.' COVID-19's economic toll hits La
10	2020-11-17T07:44:34Z	Elsevier BV	10.1016/j.outlook.2020.08.013	journal-article	Crossref	Aging in America: How COVID-19 Will Change Care, Coverage, and Compassion
13	2020-07-02T14:40:29Z	Middle Atlantic Review of Latin America	10.23870/marlas.312	journal-article	Crossref	América Latina y el Covid-19
14	2020-12-21T22:17:22Z	Editorial Universidad de Sevilla	10.12795/araucaria.2020.i45.16	journal-article	Crossref	América Latina y la Unión Europea: agendas sociales, competencia geopolítica y t
15	2020-12-08T05:21:13Z	Universidad Nacional Autónoma de Me	10.22201/fe.18701442e.2020.37.77770	journal-article	Crossref	América Latina: Hacia un período de débil crecimiento y COVID-19
16	2021-01-15T21:23:47Z	Elsevier BV	10.1016/j.explore.2020.08.012	journal-article	Crossref	America, consciousness, COVID-19, climate change, and migration
20	2021-03-29T14:24:45Z	Oxford University Press (OUP)	10.1093/jtm/taaa176	journal-article	Crossref	Analysis of travel restrictions for COVID-19 control in Latin America through netw
22	2020-07-26T13:40:28Z	Project Muse	10.1353/lag.2020.0049	journal-article	Crossref	As organizações religiosas brasileiras frente à pandemia de Covid-19
23	2021-01-25T05:05:46Z	Project Muse	10.1353/lag.2020.0057	journal-article	Crossref	Asylum and Mass Detention at the U.S.-Mexico Border during Covid-19
24	2020-09-10T02:11:21Z	Human Kinetics	10.1123/ijsc.2020-0217	journal-article	Crossref	Australian Football in America During COVID-19
26	2021-05-09T11:41:00Z	Ovid Technologies (Wolters Kluwer Heal	10.1097/ede.0000000000001293	journal-article	Crossref	Being a Latin American Woman in Science During the COVID-19 Pandemic
27	2020-11-16T14:47:30Z	Clinical Biotec	10.21931/rb/2020.05.04.27	journal-article	Crossref	Bioethical Guidelines of 'Extreme Triage' Under Covid: The Question of 'Possible I
30	2020-10-29T16:24:57Z	BMJ	10.1136/bmjhci-2020-100159	journal-article	Crossref	Collaboration in times of COVID-19: the urgent need for open-data sharing in Lati
34	2020-09-02T12:26:39Z	SAGE Publications	10.1177/0003134820927313	journal-article	Crossref	Coronavirus Disease 2019 (COVID-19) and Surgical Recommendations in Latin Am
35	2020-10-16T08:43:43Z	Gavin Publishers	10.29011/2690-9480.100122	journal-article	Crossref	Corporate Responses to COVID-19 Layoffs in North America and the Role of Hum
36	2020-08-03T07:17:11Z	International Institute for Science, Tech	10.7176/jlpg/99-13	journal-article	Crossref	COVID -19 Emergency Laws and Law Enforcement in Nigeria America And Britain.
37	2021-02-22T15:43:20Z	Informa UK Limited	10.1080/07399332.2020.1833884	journal-article	Crossref	COVID 19: sexual vulnerabilities and gender perspectives in Latin America
38	2021-02-08T21:16:49Z	Wiley	10.1111/blar.13180	journal-article	Crossref	COVID -19 and Historical Global Rupture in Latin America
39	2021-02-08T20:54:36Z	Wiley	10.1111/blar.13188	journal-article	Crossref	COVID -19 and the Limitations of Official Responses to Gender-Based Violence in

Selección de URIS

Para el uso de los prefijos para la generación de las tripletas se basaron en las siguientes:

Prefijo	URI	Tipo
dataPrefix	http://ky.utpl.edu.ec/publicicovid/data	Definida
EventPrefix	http://purl.org/NET/c4dm/event.owl	Establecida en la web
CPrefix	http://purl.org/spar/c4o/	Establecida en la web
vcard	http://www.w3.org/2006/vcard/ns	Establecida en la web
foaf	http://xmlns.com/foaf/0.1/	Establecida en la web
dbo	http://dbpedia.org/ontology/	Establecida en la web
vivo	http://vivoweb.org/ontology/core	Establecida en la web
bibo	http://purl.org/ontology/bibo	Establecida en la web
dct	http://purl.org/dc/terms/	Establecida en la web
rdf	http://www.w3.org/1999/02/22-rdf-syntax-ns	Establecida en la web

Generación Jena

En esta sección se explicará la creación de las Tripletas-RDF que se toma como base los datos del archivo CSV con la finalidad de generar un nuevo archivo con extensión rdf. Este proceso se lo realizó mediante el lenguaje de programación Java con el IDE “Apache NetBeans IDE 12” y la librería Jena que proporciona todas las funcionalidades necesarias. Además, se realizó en enriquecimiento de los datos del archivo CSV mediante el lenguaje Python gracias a la API de “tagme” que permite la información de nombre de organizaciones mediante texto.

El primer paso fue realizar consultas a la Api de tagme para obtener algunos datos de las publicaciones o investigaciones. Este proceso tiene como resultado buscar si un paper o investigación es afiliado a una organización, el texto se lo obtuvo del mismo archivo .csv donde se encuentra las publicaciones y este se manda a tagme donde nos retornara las respuestas mediante un REQUEST (Ver la siguiente figura).

```

Archivo Edición Formato Ver Ayuda
["101055a-1229-5048", "http://dbpedia.org/resource/ISWC", 0.24496769905090332, 0.31707316637039185, "http://dbpedia.org/resource/Category:Music_technology"]
["101055a-1229-5048", "http://dbpedia.org/resource/ISWC", 0.24496769905090332, 0.31707316637039185, "http://dbpedia.org/resource/Category:ISO_standards"]
["101055a-1229-5048", "http://dbpedia.org/resource/ISWC", 0.24496769905090332, 0.31707316637039185, "http://dbpedia.org/resource/Category:Identifiers"]
["101055a-1229-5048", "http://dbpedia.org/resource/ISWC", 0.24496769905090332, 0.31707316637039185, "http://dbpedia.org/resource/Category:Universal_identifiers"]
["101055a-1229-5048", "http://dbpedia.org/resource/RDF", 0.3218223750591278, 0.4697193503379822, "http://dbpedia.org/resource/Category:Resource_Description_Framework"]
["101055a-1229-5048", "http://dbpedia.org/resource/RDF", 0.3218223750591278, 0.4697193503379822, "http://dbpedia.org/resource/Category:Knowledge_representation"]
["101055a-1229-5048", "http://dbpedia.org/resource/RDF", 0.3218223750591278, 0.4697193503379822, "http://dbpedia.org/resource/Category:World_Wide_Web_Consortium_standards"]
["101055a-1229-5048", "http://dbpedia.org/resource/RDF", 0.3218223750591278, 0.4697193503379822, "http://dbpedia.org/resource/Category:XML"]
["101055a-1229-5048", "http://dbpedia.org/resource/RDF", 0.3218223750591278, 0.4697193503379822, "http://dbpedia.org/resource/Category:XML-based_standards"]
["101055a-1229-5048", "http://dbpedia.org/resource/RDF", 0.3218223750591278, 0.4697193503379822, "http://dbpedia.org/resource/Category:Metadata"]
["101055a-1229-5048", "http://dbpedia.org/resource/RDF", 0.3218223750591278, 0.4697193503379822, "http://dbpedia.org/resource/Category:Semantic_Web"]
["101055a-1229-5048", "http://dbpedia.org/resource/RDF", 0.3218223750591278, 0.4697193503379822, "http://dbpedia.org/resource/Category:Bibliography_file_formats"]
["101055a-1229-5048", "http://dbpedia.org/resource/RDF_triples", 0.19807380437850952, 0.222222238779068, "http://dbpedia.org/resource/Category:Resource_Description_Framework"]
["101055a-1229-5048", "http://dbpedia.org/resource/RDF_triples", 0.19807380437850952, 0.222222238779068, "http://dbpedia.org/resource/Category:Knowledge_representation"]
["101055a-1229-5048", "http://dbpedia.org/resource/RDF_triples", 0.19807380437850952, 0.222222238779068, "http://dbpedia.org/resource/Category:World_Wide_Web_Consortium_standards"]
["101055a-1229-5048", "http://dbpedia.org/resource/RDF_triples", 0.19807380437850952, 0.222222238779068, "http://dbpedia.org/resource/Category:XML"]
["101055a-1229-5048", "http://dbpedia.org/resource/RDF_triples", 0.19807380437850952, 0.222222238779068, "http://dbpedia.org/resource/Category:XML-based_standards"]
["101055a-1229-5048", "http://dbpedia.org/resource/RDF_triples", 0.19807380437850952, 0.222222238779068, "http://dbpedia.org/resource/Category:Metadata"]
["101055a-1229-5048", "http://dbpedia.org/resource/RDF_triples", 0.19807380437850952, 0.222222238779068, "http://dbpedia.org/resource/Category:Semantic_Web"]
["101055a-1229-5048", "http://dbpedia.org/resource/RDF_triples", 0.19807380437850952, 0.222222238779068, "http://dbpedia.org/resource/Category:Bibliography_file_formats"]
["101055a-1229-5048", "http://dbpedia.org/resource/triples", 0.17579340934753418, 0.35158681869506836, "http://dbpedia.org/resource/Category:Baseball_terminology"]
["101055a-1229-5048", "http://dbpedia.org/resource/triples", 0.17579340934753418, 0.35158681869506836, "http://dbpedia.org/resource/Category:Baseball_statistics"]
["101055a-1229-5048", "http://dbpedia.org/resource/knowledge_graph", 0.632880449295044, 0.9714285731315613, "http://dbpedia.org/resource/Category:Knowledge_engineering"]
["101055a-1229-5048", "http://dbpedia.org/resource/knowledge_graph", 0.632880449295044, 0.9714285731315613, "http://dbpedia.org/resource/Category:Technical_communication"]
["101055a-1229-5048", "http://dbpedia.org/resource/knowledge_graph", 0.632880449295044, 0.9714285731315613, "http://dbpedia.org/resource/Category:Information_science"]
["101055a-1229-5048", "http://dbpedia.org/resource/knowledge_graph", 0.632880449295044, 0.9714285731315613, "http://dbpedia.org/resource/Category:Semantic_Web"]
["101055a-1229-5048", "http://dbpedia.org/resource/knowledge_graph", 0.632880449295044, 0.9714285731315613, "http://dbpedia.org/resource/Category:Ontology_(information_science)"]
["101055a-1229-5048", "http://dbpedia.org/resource/knowledge_graph", 0.632880449295044, 0.9714285731315613, "http://dbpedia.org/resource/Category:Knowledge_representation"]
["101055a-1229-5048", "http://dbpedia.org/resource/DrugBank", 0.5851426887512207, 1, "http://dbpedia.org/resource/Category:Chemical_databases"]
["101055a-1229-5048", "http://dbpedia.org/resource/DrugBank", 0.5851426887512207, 1, "http://dbpedia.org/resource/Category:Metabolic_databases"]
["101055a-1229-5048", "http://dbpedia.org/resource/GERBIL", 0.31589147448539734, 0.6317829489707947, "http://dbpedia.org/resource/Category:Gerbils"]
["101055a-1229-5048", "http://dbpedia.org/resource/GERBIL", 0.31589147448539734, 0.6317829489707947, "http://dbpedia.org/resource/Category:Pet_rodents"]
["101055a-1229-5048", "http://dbpedia.org/resource/GERBIL", 0.31589147448539734, 0.6317829489707947, "http://dbpedia.org/resource/Category:Fauna_of_Central_Asia"]
["101055a-1229-5048", "http://dbpedia.org/resource/GERBIL", 0.31589147448539734, 0.6317829489707947, "http://dbpedia.org/resource/Category:Fauna_of_Mongolia"]
["101055a-1229-5048", "http://dbpedia.org/resource/ISWC", 0.2306729555130005, 0.31707316637039185, "http://dbpedia.org/resource/Category:Music_technology"]
["101055a-1229-5048", "http://dbpedia.org/resource/ISWC", 0.2306729555130005, 0.31707316637039185, "http://dbpedia.org/resource/Category:ISO_standards"]
["101055a-1229-5048", "http://dbpedia.org/resource/ISWC", 0.2306729555130005, 0.31707316637039185, "http://dbpedia.org/resource/Category:Identifiers"]
["101055a-1229-5048", "http://dbpedia.org/resource/ISWC", 0.2306729555130005, 0.31707316637039185, "http://dbpedia.org/resource/Category:Universal_identifiers"]

```

Esta nueva información será generada en un archivo txt, para después extraer los datos y acoplarlos al archivo original de las investigaciones.

Una vez obtenido todos los datos necesarios se procedió a generar el archivo rdf con la librería Jena. El proceso para la generación de datos RDF consiste en primer lugar establecer los prefijos a un modelo, ver la siguiente figura.

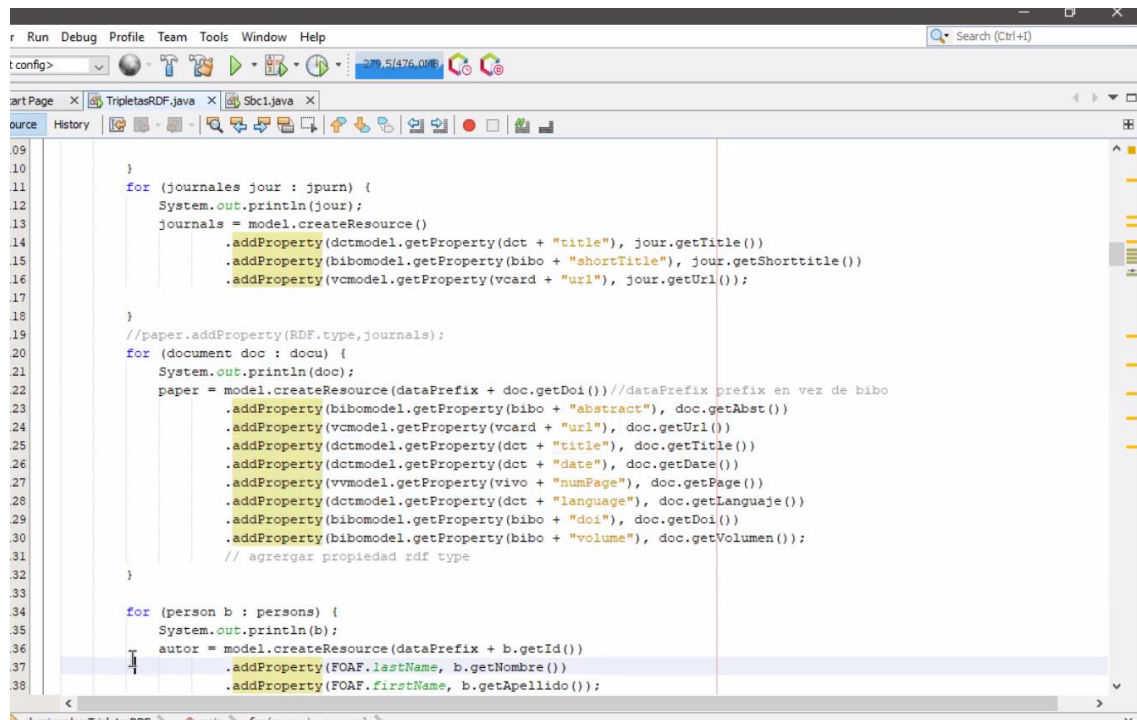
```

1  Run Debug Profile Team Tools Window Help
2  Search (Ctrl+F)
3  ult config>
4  Start Page X TripletasRDF.java X Sbci.java X
5  Source History
6  73      model.setNsPrefix("c4o", CPrefix);
7  74      Model cpmodel = ModelFactory.createDefaultModel();
7  75
7  76      String vcard = "http://www.w3.org/2006/vcard/ns#";
7  77      model.setNsPrefix("vcard", vcard);
7  78      Model vcmodel = ModelFactory.createDefaultModel();
7  79
7  80      String foaf = "http://xmlns.com/foaf/0.1/";
7  81      model.setNsPrefix("foaf", foaf);
7  82      Model foafmodel = ModelFactory.createDefaultModel();
7  83
7  84      String dbo = "http://dbpedia.org/ontology/";
7  85      model.setNsPrefix("dbo", dbo);
7  86      Model dboModel = ModelFactory.createDefaultModel();
7  87
7  88      String vivo = "http://vivoweb.org/ontology/core#";
7  89      model.setNsPrefix("vivo", vivo);
7  90      Model vmmodel = ModelFactory.createDefaultModel();
7  91
7  92      String bibo = "http://purl.org/ontology/bibo#";
7  93      model.setNsPrefix("bibo", bibo);
7  94      Model bibomodel = ModelFactory.createDefaultModel();
7  95
7  96      String dct = "http://purl.org/dc/terms/";
7  97      model.setNsPrefix("dct", dct);
7  98      Model dctmodel = ModelFactory.createDefaultModel();
7  99
7  100      String rdf = "http://www.w3.org/1999/02/22-rdf-syntax-ns#";
7  101      model.setNsPrefix("rdf", rdf);
7  102      Model rdfmodel = ModelFactory.createDefaultModel();

```

Después se realizó una lectura de los datos del archivo csv donde se encuentran las publicaciones y esta lectura almacena los datos en un ArrayList, estos son leídos y almacenados en clases creadas para tener un mejor control de los campos de las

publicaciones al momento de crear el rdf. Estos datos son obtenidos por medio de una estructura “for” para generar las triples que se pueden apreciar en la siguiente imagen.



```
09
10
11     for (journals jour : jpurn) {
12         System.out.println(jour);
13         journals = model.createResource()
14             .addProperty(dctmodel.getProperty(dct + "title"), jour.getTitle())
15             .addProperty(bibomodel.getProperty(bibo + "shortTitle"), jour.getShorttitle())
16             .addProperty(vcardmodel.getProperty(vcard + "url"), jour.getUrl());
17     }
18
19     //paper.addProperty(RDF.type, journals);
20     for (document doc : docu) {
21         System.out.println(doc);
22         paper = model.createResource(dataPrefix + doc.getDoi()); //dataPrefix prefix en vez de bibo
23         .addProperty(bibomodel.getProperty(bibo + "abstract"), doc.getAbst())
24         .addProperty(vcardmodel.getProperty(vcard + "url"), doc.getUrl())
25         .addProperty(dctmodel.getProperty(dct + "title"), doc.getTitle())
26         .addProperty(dctmodel.getProperty(dct + "date"), doc.getDate())
27         .addProperty(vvmodel.getProperty(vivo + "numPage"), doc.getPage())
28         .addProperty(dctmodel.getProperty(dct + "language"), doc.getLanguaje())
29         .addProperty(bibomodel.getProperty(bibo + "doi"), doc.getDoi())
30         .addProperty(bibomodel.getProperty(bibo + "volume"), doc.getVolumen());
31         // agregar propiedad rdf type
32     }
33
34     for (person b : persons) {
35         System.out.println(b);
36         autor = model.createResource(dataPrefix + b.getId())
37             .addProperty(FOAF.lastName, b.getNombre())
38             .addProperty(FOAF.firstName, b.getApellido());
```

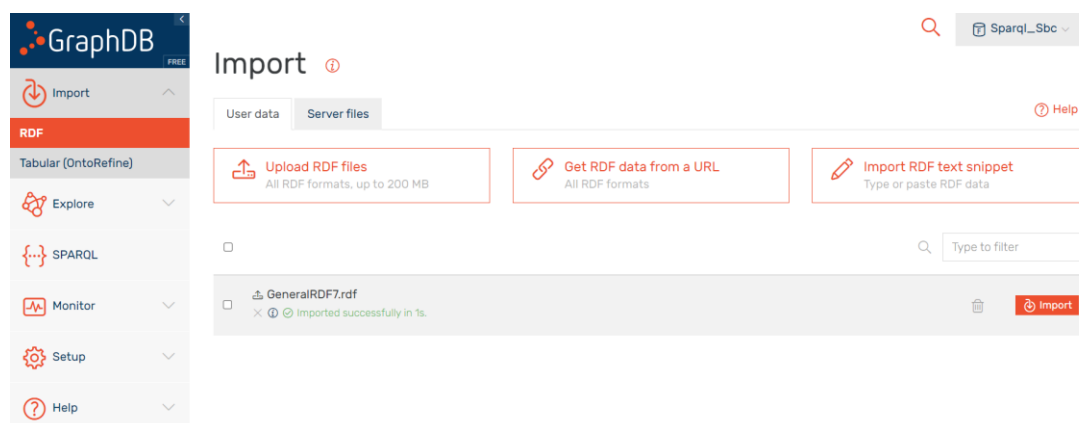
En la siguiente figura se puede apreciar un extracto sobre las tripletas generadas por Jena.

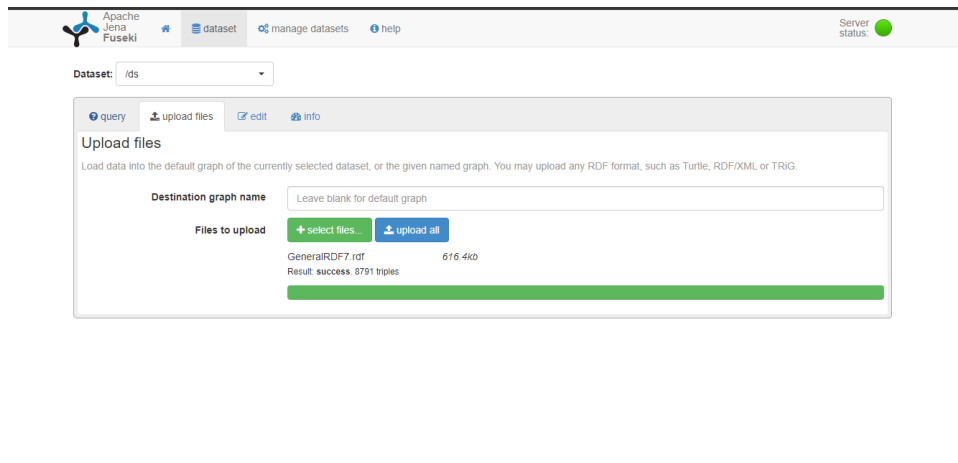
```
GeneralRDF4: Bloc de notas
Archivo Edición Formato Ver Ayuda

<rdf:RDF
  xmlns:dct="http://purl.org/dc/terms/"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:c4o="http://purl.org/spar/c4o/"
  xmlns:event="http://purl.org/NET/c4dm/event.owl#"
  xmlns:dbo="http://dbpedia.org/ontology/"
  xmlns:vcard="http://www.w3.org/2006/vcard/ns#"
  xmlns:data="http://ky.utpl.edu.ec/publicicovid/data#"
  xmlns:bibo="http://purl.org/ontology/bibo#"
  xmlns:vivo="http://vivoweb.org/ontology/core#"
  xmlns:foaf="http://xmlns.com/foaf/0.1/">
  <rdf:Description>
    <vcard:url>N/A</vcard:url>
    <bibo:shortTitle>N/A</bibo:shortTitle>
    <dct:title>&lt;strong&gt;&lt;/strong&gt;COVID-19: Prediction of Vulnerable Areas Having High
  </rdf:Description>
  <rdf:Description>
    <vcard:url>N/A</vcard:url>
    <bibo:shortTitle>N/A</bibo:shortTitle>
    <dct:title>Drivers and Barriers to Living in a Multigenerational Household Pre-COVID - Mid-
  </rdf:Description>
  <rdf:Description>
    <vcard:url>N/A</vcard:url>
    <bibo:shortTitle>Social Work and the COVID-19 Pandemic</bibo:shortTitle>
    <dct:title>COVID-19 and Social Inequalities: A Political View From Social Work</dct:title>
  </rdf:Description>
  <rdf:Description rdf:about="http://ky.utpl.edu.ec/publicicovid/data#191">
    <foaf:lastName>Thiago</foaf:lastName>
    <vcard:url>http://dx.doi.org/10.4324/9781003154037-19</vcard:url>
    <bibo:doi>191</bibo:doi>
    <data:doiDocu>1043249781003154037-21</data:doiDocu>
    <data:idpersona>191</data:idpersona>
    <bibo:volume>N/A</bibo:volume>
    <dct:title>The US</dct:title>
    <vivo:numPage> and Covid-19</vivo:numPage>
    <dct:language>5/4/2021</dct:language>
```

Almacenamiento de los datos RDF

Antes de poder usar estos datos estructurados en alguna aplicación es necesario subirlos en una base de datos para tener un control mucho más gestionado sobre la información de las publicaciones, por ello se usó la base de grafos denominada “GraphDB” y “Apache Jena fuseki”.





Consultas SPARQL

En si se ha creado apis en el servidor en donde cada una genera consultas hacia el repositorio semántico. Y con la información que se obtiene se procede a generar gráficas y tablas estadísticas en las interfaces de usuario. Las consultas que se han generado a partir de los datos que se encuentra en GrahpDB y Apache Jena Fuseki fueron las siguientes:

Pregunta 1: ¿Cuáles son las publicaciones en ingles y sus autores?

Consulta:

```
pub_Autor X ⊕
8 PREFIX dc: <http://purl.org/dc/elements/1.1/>
9 PREFIX dct: <http://purl.org/dc/terms/>
10 PREFIX data: <http://ky.utpl.edu.ec/publicicovid/data/>
11 #Publicaciones en Ingles con el Nombre del Autor
12 SELECT DISTINCT ?nombreDocumento ?doi ?nombre WHERE {
13     ?documento dct:title ?nombreDocumento.
14     ?documento dct:language "en".
15     ?documento bibo:doi ?doi .
16     ?autorid data:doiDocu ?doi .
17     ?autorid data:idpersona ?id .
18     ?idpersona data:idpersona ?id .
19     ?idpersona foaf:firstName ?nombre
20 }LIMIT 100
21
```


Resultado:

	nombreDocumento	doi	nombre
1	"A Note on the Potential BCG Vaccination ? COVID-19 Molecular Link"	"1021742666796701999200629003417"	"Alex Y."
2	"A Note on the Potential BCG Vaccination ? COVID-19 Molecular Link"	"1021742666796701999200629003417"	"Piotr"
3	"A Note on the Potential BCG Vaccination ? COVID-19 Molecular Link"	"1021742666796701999200629003417"	"Alex"
4	"Education in COVID-19 and Post-COVID Eras: A Paradigm Shift"	"102139ssrn3827633"	"Samuel A."
5	"Education in COVID-19 and Post-COVID Eras: A Paradigm Shift"	"102139ssrn3827633"	"Oluwatosin F."
6	"Education in COVID-19 and Post-COVID Eras: A Paradigm Shift"	"102139ssrn3827633"	"Ikeoluwapo"
7	"Psychological Impact of Coronavirus (COVID-19) Disease on Cancer Patients"	"1021742666796701999210101151035"	"Gaurav"
8	"Potential Drug Options for Treatment of COVID-19: A Review"	"1021742666796701999200701131604"	"Dharmendra"

Pregunta 2: ¿Cuáles son todas las publicaciones y fechas de publicación?

pub_Autor	pub_Date	
5	PREFIX tab: <http://purl.org/tab/ns#>	
6	PREFIX bib: <http://zeitkunst.org/bibtex/0.1/bibtex.owl#>	
7	PREFIX co: <http://purl.org/ontology/co/core#>	
8	PREFIX dc: <http://purl.org/dc/elements/1.1/>	
9	PREFIX dct: <http://purl.org/dc/terms/>	
10	PREFIX data: <http://ky.utpl.edu.ec/publicicovid/data/>	
11	PREFIX fabio: <http://purl.org/spar/fabio/>	
12	#Todas las publicaciones, nombre y fecha de publicación	
13	SELECT DISTINCT ?doi ?date ?nombreDocumento where {	
14	?documento dct:title ?nombreDocumento.	
15	?documento bibo:doi ?doi .	
16	?documento dct:date ?date	
17	}	
18	LIMIT 100	

Resultado:

	doi	date	nombreDoc 
1	"10.2307/j.ctv1850gc4.23"	"21/10/2020"	"Roma communities' experiences of racism during the COVID-19 pandemic"
2	"10.46692/9781447360377.021"	"8/4/2021"	"Rethinking Right and Wrong: Social Work, COVID-19 and the Crisis of Ethics"
3	"10.2307/j.ctv18gfz7c.20"	"25/10/2020"	"The Economics of COVID-19 in Italy and Lessons for Africa"
4	"10.2307/j.ctv1850gc4.16"	"21/10/2020"	"The extent of the COVID-19 crisis in South Korea"
5	"10.2174/266679670199920116212443"	"20/11/2020"	"COVID-19 Effect: Time to Change the Way of Research"
6	"10.1016/b978-0-323-82860-4.00021-5"	"4/6/2021"	"The disproportionate burden of COVID-19 in Africa"
7	"10.1136/bmj.m4386"	"11/11/2020"	"Covid-19: Nearly 20% of patients receive psychiatric diagnosis within three months of covid, study finds"
8	"10.36104/amc.2020.2052"	"21/2/2021"	"COVID-19 y SARS-CoV-2"
9	"10.51432/978-1-8381524-8-2_7"	"21/5/2021"	"Complex Risk Analysis of E-Commerce Companies Related to COVID 19"
10	"10.2307/j.ctv1850gc4.17"	"21/10/2020"	"Social work and COVID-19 in Greece"

Pregunta 3: ¿Cuáles son los nombres de las publicaciones, los títulos cortos y el url?

```
pub_Autor@en X pub_Date X pub_ShortTitle X pub_Autor@es X pub_NroDoc_Autores X
5 PREFIX fab: <http://purl.org/fab/ns#>
6 PREFIX bib: <http://zeitkunst.org/bibtex/0.1/bibtex.owl#>
7 PREFIX co: <http://purl.org/ontology/co/core#>
8 PREFIX dc: <http://purl.org/dc/elements/1.1/>
9 PREFIX dct: <http://purl.org/dc/terms/>
10 PREFIX data: <http://ky.utpl.edu.ec/publicicovid/data/>
11 PREFIX fabio: <http://purl.org/spar/fabio/>
12 PREFIX vcard: <http://www.w3.org/2006/vcard/ns#>
13 #Publicaciones Titulo, titulo corto y URL
14 SELECT DISTINCT ?nombreDocumento ?shortDocumento ?url WHERE {
15     ?documento dct:title ?nombreDocumento.
16     ?documento bibo:shortTitle ?shortDocumento.
17     ?url vcard:url ?shortDocumento
18 }
```

Resultado:

	nombreDocumento	shortDocumento	url
1	"The importance of publishing non-COVID-19 rese arch during COVID-19"	"Journal of Clinical Epidemiology"	data:10.2307/j.ctv1850gc4.23
2	"The importance of publishing non-COVID-19 rese arch during COVID-19"	"Journal of Clinical Epidemiology"	_:genid-d49528daa87c430e81c976dcf4fbd5d7-B 57CA67928586DA264646EE230366DCF
3	"The importance of publishing non-COVID-19 rese arch during COVID-19"	"Journal of Clinical Epidemiology"	_:genid-d49528daa87c430e81c976dcf4fbd5d7-9 AA37243ADB9D6E045EB1FF2AAAF59C1
4	"The importance of publishing non-COVID-19 rese arch during COVID-19"	"Journal of Clinical Epidemiology"	_:genid-d49528daa87c430e81c976dcf4fbd5d7-n ode123
5	"The importance of publishing non-COVID-19 rese arch during COVID-19"	"Journal of Clinical Epidemiology"	data:10.46692/9781447360377.021
6	"The importance of publishing non-COVID-19 rese arch during COVID-19"	"Journal of Clinical Epidemiology"	_:genid-d49528daa87c430e81c976dcf4fbd5d7-n ode124
7	"The importance of publishing non-COVID-19 rese arch during COVID-19"	"Journal of Clinical Epidemiology"	data:10.2307/j.ctv18gfz7c.20

Pregunta 4: ¿Cuáles son las publicaciones en español y sus autores?

```
pub_Autor@en X pub_Date X pub_ShortTitle X pub_Autor@es X pub_NroDoc_Autores X
7 PREFIX co: <http://purl.org/ontology/co/core#>
8 PREFIX dc: <http://purl.org/dc/elements/1.1/>
9 PREFIX dct: <http://purl.org/dc/terms/>
10 PREFIX data: <http://ky.utpl.edu.ec/publicicovid/data/>
11 #Publicaciones en Español con el Nombre del Autor
12 SELECT DISTINCT ?nombreDocumento ?doi ?nombre WHERE {
13     ?documento dct:title ?nombreDocumento.
14     ?documento dct:language "es".
15     ?documento bibo:doi ?doi .
16     ?autorid data:doiDocu ?doi .
17     ?autorid data:idpersona ?id .
18     ?idpersona data:idpersona ?id .
19     ?idpersona foaf:firstName ?nombre
20 }LIMIT 1000
```

Resultado:

	nombreDocumento	doi	nombre
1	"Medical literature and COVID-19"	"103735120203122"	"JM"
2	"Fecaloid peritonitis. Positive COVID postoperative ly."	"1037351202031410"	"J"
3	"Genetic and Covid-19"	"1032440ar202013703ed01"	"Jose Miguel"
4	"Las industrias creativas versus COVID-19"	"1046840ec20201307"	"Guilherme"
5	"Thrombotic coagulopathy and COVID-19"	"1032440ar202013702doc01"	"Vicente"
6	"La ASAC en la pandemia COVID-19"	"103735120203141"	"F"

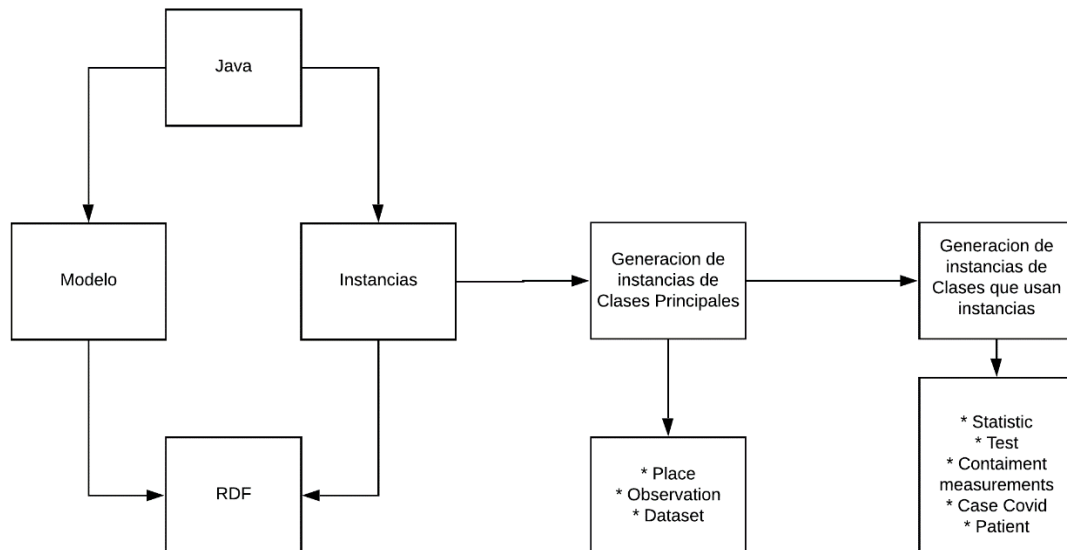
Pregunta 5: ¿Cuáles son los números de Documentos que han escrito los autores?

pub_Autor@en	pub_Date	pub_ShortTitle	pub_Autor@es	pub_NroDoc_Autores
<pre>1 / PREFIX co: <http://purl.org/ontology/co/core#> 2 PREFIX dc: <http://purl.org/dc/elements/1.1/> 3 PREFIX dct: <http://purl.org/dc/terms/> 4 PREFIX data: <http://ky.utpl.edu.ec/publicicovid/data/> 5 #Nro. de Documentos que han escrito los autores 6 7 SELECT DISTINCT ?nombreDocumento ?nombre (COUNT(?nombreDocumento) AS ?total) WHERE { 8 ?documento dct:title ?nombreDocumento. 9 ?documento bibo:doi ?doi . 10 ?autorid data:doiDocu ?doi . 11 ?autorid data:idpersona ?id . 12 ?idpersona data:idpersona ?id . 13 ?idpersona foaf:firstName ?nombre . 14 }GROUP BY ?nombreDocumento ?nombre ORDER BY DESC(?total) 15 16 17 18 19 20</pre>				

Resultado:

	nombreDocumento	nombre	total
1	"WTP for COVID-19 Job Amenities and the Effects of COVID-19 on Labor Supply"	"Helen"	"3""xsd:integer
2	"Iran and COVID-19"	"Ali"	"2""xsd:integer
3	"Infodemiological Study on COVID-19 Epidemic and COVID-19 Infodemic"	"An"	"2""xsd:integer
4	"COVID-19 and Global Inequality"	"Chrispin"	"2""xsd:integer
5	"Risks of early mortality and pulmonary complications following surgery in patients with COVID-19"	"E"	"2""xsd:integer
6	"COVID-19 and Global Inequality"	"Gerard"	"2""xsd:integer
7	"The Economics of COVID-19 in Italy and Lessons for Africa"	"Giovanni"	"2""xsd:integer
8	"Africa and the Economic Pathologies of the COVID-19 Pandemic"	"Howard"	"2""xsd:integer
9	"COVID-19"	"J. Michael"	"2""xsd:integer

Funcionamiento de la generación de datos RDF



Aplicación:

Para la aplicación se ha tomado en cuenta un pequeño código Python que permite consultar las sentencias Sparql antes descritas, el cual mediante cada consulta nos devolverá la información almacenada en la base de datos.

Se muestra a continuación:

```

"C:\Users\Alexis Montoya\AppData\Local\Programs\Python\Python37\python.exe" "C:
1. ¿Cuáles son las publicaciones en ingles y sus autores?
2. ¿Cuáles son todas las publicaciones y fechas de publicación?
3. ¿Cuáles son los nombres de las publicaciones, los títulos cortos y el url?
4. ¿Cuáles son las publicaciones en español y sus autores?
5. ¿Cuáles son los números de Documentos que han escrito los autores?
  
```

Para el consumo del endpoint se utiliza lo siguiente en cada una de las preguntas:

```

sparql = SPARQLWrapper(
    'http://localhost:3030/ds/sparql')
def get_1():
    sparql.setQuery('''
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
PREFIX vivo: <http://vivoweb.org/ontology/core#>
PREFIX bibo: <http://purl.org/ontology/bibo#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX fab: <http://purl.org/fab/ns#>
PREFIX bib: <http://zeitkunst.org/bibtex/0.1/bibtex.owl#>
PREFIX co: <http://purl.org/ontology/co/core#>
PREFIX dc: <http://purl.org/dc/elements/1.1/>
PREFIX dct: <http://purl.org/dc/terms/>
PREFIX data: <http://ky.utpl.edu.ec/publicicovid/data/>

#Publicaciones en Ingles con su apellido
select DISTINCT ?nombreDocumento ?doi ?apellido where {
?documento dct:title ?nombreDocumento.
?documento dct:language "en".
?documento bibo:doi ?doi .
?autorid data:doiDocu ?doi .
?autorid data:idpersona ?id .
?idpersona data:idpersona ?id .
?idpersona foaf:firstName ?apellido
}limit 100

''')
sparql.setReturnFormat(JSON)

```

Resultado:

```

nombreDocumento: SEAMHCRD deterministic compartmental model based on clinical stages of infection for COVID-19 pandemic in Sultanate of Oman doi: 101038s41598
nombreDocumento: SEAMHCRD deterministic compartmental model based on clinical stages of infection for COVID-19 pandemic in Sultanate of Oman doi: 101038s41598
nombreDocumento: SEAMHCRD deterministic compartmental model based on clinical stages of infection for COVID-19 pandemic in Sultanate of Oman doi: 101038s41598
nombreDocumento: SEAMHCRD deterministic compartmental model based on clinical stages of infection for COVID-19 pandemic in Sultanate of Oman doi: 101038s41598
nombreDocumento: SEAMHCRD deterministic compartmental model based on clinical stages of infection for COVID-19 pandemic in Sultanate of Oman doi: 101038s41598
nombreDocumento: SEAMHCRD deterministic compartmental model based on clinical stages of infection for COVID-19 pandemic in Sultanate of Oman doi: 101038s41598
nombreDocumento: SEAMHCRD deterministic compartmental model based on clinical stages of infection for COVID-19 pandemic in Sultanate of Oman doi: 101038s41598
nombreDocumento: COVID-19: A daily dose of yoga to cope with anxiety doi: 1021742666796701999201123102216 apellido:Rima
nombreDocumento: COVID-19: A daily dose of yoga to cope with anxiety doi: 1021742666796701999201123102216 apellido:Surabhi
nombreDocumento: Higher Education System in Light of Covid-19 doi: 1034172ps202045 apellido:Mohammad Reza
nombreDocumento: Higher Education System in Light of Covid-19 doi: 1034172ps202045 apellido:Jalal

```

Enlace Github: <https://github.com/Jtaramon/ProyectoSbc>