



UNIVERSIDAD TÉCNICA PARTICULAR DE LOJA

La Universidad Católica de Loja

AREA TÉCNICA

**TITULACIÓN DE INGENIERO EN SISTEMAS INFORMÁTICOS Y
COMPUTACIÓN**

**Publicación de datos enlazados observando los principios de Linked
Data en el ámbito de Legislación de Educación Superior**

TRABAJO DE FIN DE TITULACIÓN

AUTOR: Valverde Vásquez, María Fernanda

DIRECTOR: Piedra Pullaguari, Nelson Oswaldo, Ing.

LOJA-ECUADOR

2014

APROBACIÓN DEL DIRECTOR DEL TRABAJO DE FIN DE TITULACIÓN

Ing.

Nelson Oswaldo Piedra Pullaguari.

DIRECTOR DEL TRABAJO DE FIN DE TITULACIÓN

De mi consideración:

El presente trabajo de fin de titulación: **“Publicación de datos enlazados observando los principios de Linked Data en el ámbito de Legislación de Educación Superior”** realizado por **María Fernanda Valverde Vásquez**, ha sido orientado y revisado durante su ejecución, por eso se aprueba la presentación del mismo.

Loja, Abril de 2014

f) -----

DECLARACIÓN DE AUTORÍA Y CESIÓN DE DERECHOS

“Yo **María Fernanda Valverde Vásquez** declaro ser autor (a) del presente trabajo de fin de titulación: “**Publicación de datos enlazados observando los principios de Linked Data en el ámbito de Legislación de Educación Superior**”, de la **Titulación de Ingeniero en Sistemas Informáticos y Computación**, siendo **Nelson Oswaldo Piedra Pullaguari** director (a) del presente trabajo; y eximo expresamente a la Universidad Técnica Particular de Loja y a sus representantes legales de posibles reclamos o acciones legales. Además certifico que las ideas, conceptos, procedimientos y resultados vertidos en el presente trabajo investigativo, son de mi exclusiva responsabilidad.

Adicionalmente declaro conocer y aceptar la disposición del Art. 67 del Estatuto Orgánico de la Universidad Técnica Particular de Loja que en su parte pertinente textualmente dice: “Forman parte del patrimonio de la Universidad la propiedad intelectual de investigaciones, trabajos científicos o técnicos y tesis de grado que se realicen a través, o con el apoyo financiero, académico o institucional (operativo) de la Universidad”

f) -----

Autor: María Fernanda Valverde Vásquez

Cedula: 1104724149

DEDICATORIA

A:

Dios, por darme la oportunidad de vivir y por estar conmigo en cada paso que doy, por fortalecer mi corazón e iluminar mi mente y por haber puesto en mi camino a aquellas personas que han sido mi soporte y compañía durante todo el periodo de estudio.

A mis padres Augusto Valverde y María Eugenia Vásquez, por darme la vida, quererme mucho, creer en mí y porque siempre me apoyaron y supieron guiarme por el camino correcto, por esto y mucho más les dedico este meta alcanzada.

A mis abuelitos Luis Alberto y Herlinda Sarango (Mamita Hila), por quererme y apoyarme siempre, esto también se lo debo a ustedes.

A mi tía Carmen mi segunda madre por ser mi ejemplo a seguir, por su apoyo y cariño incondicional que desde pequeña me brindo. Gracias tía querida.

Mis hermanos Gabriela Alexandra, Luis Augusto, Paola Silvana y Doménica Abigail, por estar conmigo y apoyarme siempre, los quiero mucho.

A ti cariño por estar pendiente de mí a cada instante por ser mi apoyo, gracias por regalarme muchos años de alegría y felicidad.

María Fernanda Valverde Vásquez

AGRADECIMIENTO

A mis padres quienes a lo largo de toda mi vida me han apoyado y motivado mi formación personal y académica.

Agradezco a mi director de tesis Ing. Nelson Piedra, Ing. Juan Carlos Morocho, Ing. Janeth Chicaiza quienes a lo largo de este tiempo me han guiado y compartido sus conocimientos para que pueda desarrollar el proyecto de investigación.

A mis amigos por su amistad y apoyo.

A mis profesores a quienes les debo gran parte de mis conocimientos, gracias a su paciencia y enseñanza y finalmente un eterno agradecimiento a esta prestigiosa universidad la cual abrió y abre sus puertas a jóvenes como nosotros, preparándonos para un futuro competitivo y formándonos como personas de bien.

María Fernanda Valverde Vásquez

INDICE GENERAL

APROBACIÓN DEL DIRECTOR DEL TRABAJO DE FIN DE TITULACIÓN	II
DECLARACIÓN DE AUTORÍA Y CESIÓN DE DERECHOS	III
DEDICATORIA	IV
AGRADECIMIENTO	V
INDICE GENERAL	VI
INDICE DE ILUSTRACIONES	IX
INDICE DE TABLAS	X
RESUMEN EJECUTIVO	XI
ABSTRACT	XII
INTRODUCCIÓN	1
OBJETIVOS	3
PROBLEMA	4
1. ESTADO DEL ARTE	5
1.1. Introducción.	6
1.2. Evolución de la Web.	6
1.3. Web Semántica.	8
1.3.1. Concepto de Web Semántica.....	9
1.3.2. Campos de aplicación de la Web Semántica.	11
1.3.3. Descripción de cada uno de las capas de la web semántica.	12
1.3.4. Ventajas y desventajas de la Web Semántica.....	13
1.4. Linked Data.....	13
1.4.1. Principios de Linked data.	14
1.4.2. Proceso para publicación de datos enlazados.....	16
1.5. Tecnologías Semánticas.	18
1.5.1. URI.....	18
1.5.2. RDF.....	18
1.5.2.1. Elementos primordiales de RDF.	19
1.5.2.2. Sentencia RDF.....	19
1.5.2.3. Áreas de aplicación de RDF.	20
1.5.2.4. Propiedades RDF.	21
1.5.3. RDFS.	21
1.5.4. SPARQL.....	22
1.5.4.1. Sintaxis SPARQL.	23
1.6. Vocabularios.....	23

1.6.1. DUBLIN CORE.	24
1.6.2. FOAF (Friend of a Friend).....	25
1.6.3. AIISO(Academic Institution Internal Structure Ontology).	26
1.7. Repositorios de datos RDF.	27
1.7.1. AllegroGraph.....	27
1.7.2. Virtuoso.....	27
1.7.3. 4store.	28
1.7.4. Sesame.	28
1.8. Herramienta para generación de RDF: JENA.....	29
1.8.1. Características.....	29
1.9. Metodologías para la construcción de ontologías.	30
1.9.1. Metodología Methontology.....	30
1.9.2. Metodología Neón.	31
1.10. Trabajos Relacionados en el ámbito de legislación.	33
2. PUBLICACIÓN DE DATOS ENLAZADOS EN EL ÁMBITO DE	
LEGISLACIÓN UNIVERSITARIA.....	45
2.1. Modelo del Proyecto.	46
FASE 1	48
2.2. Identificación y Extracción de Datos.	48
2.2.1. Identificación de las fuentes de datos.	48
2.2.2. Extracción de los datos.	50
FASE 2	53
2.3. Desarrollo del Vocabulario.	53
FASE 3	63
2.4. Limpieza de Datos.	63
FASE 4	67
2.5. Generación de datos a RDF y RDFS.	67
2.5.1. Generación del RDFS.....	67
2.5.2. Generación de los datos a RDF.	67
2.5.2.1. Esquema de la aplicación.	68
FASE 5	70
2.6. Publicación de datos RDF.	70
2.6.1. Consultas SPARQL.	71
FASE 6	76
2.7. Enlazado de datos.	76

DISCUSIÓN.....	81
CONCLUSIONES.....	83
RECOMENDACIONES.....	85
TRABAJOS FUTUROS.....	86
GLOSARIO DE TÉRMINOS.....	87
BIBLIOGRAFÍA.....	88
ANEXOS.....	93
Anexo 1. Script de la base leyes desde Power Designer.....	93
Anexo 2. Conexión MYSQL desde JENA.....	96
Anexo 3. Código para generación de RDF.....	97
Anexo 4. Vocabulario en formato N3.....	102
Anexo 5. Formato RDF generado mediante Jena.....	107
Anexo 6. Instalación de Virtuoso.....	111
Anexo 7. Cargar archivo RDF en Virtuoso.....	113
Anexo 8. Consultas SPARQL.....	115
Anexo 9. Artículo Científico.....	119

INDICE DE ILUSTRACIONES

Ilustración 1Evolución de la Web (W3C)	8
Ilustración 2Buscador Actual.....	10
Ilustración 3Buscador Semántico	10
Ilustración 4 Arquitectura de la Web Semántica (W3C)	11
Ilustración 5 Nube de Linked Data (W3C, 2011)	14
Ilustración 6 Proceso de publicación de datos enlazados(Corcho & Gómez, 2010)	16
Ilustración 7 Grafo RDF	20
Ilustración 8 Ciclo de vida Methontology (Fernandez, Gomez, & Juristo, 1997)	31
Ilustración 9Escenarios de la Metodología Neón (Suárez, 2010).....	32
Ilustración 10Portada de legislation.gov.uk	34
Ilustración 11 Portada de Data.Gov.uk	35
Ilustración 12 Portal de Data.Gov	37
Ilustración 13 Ontología de normas del Congreso de Chile (Biblioteca del Congreso Nacional de Chile, 2011)	44
Ilustración 14 Modelo del Proyecto	47
Ilustración 15 Metadatos en DBpedia sobre la UTPL.....	49
Ilustración 16 Sitio Oficial del organismo CEAACES	50
Ilustración 17 Metadatos de la Norma de Educación Superior	49
Ilustración 18 Esquema Preliminar LUDLAW	58
Ilustración 19 Modelo de Legislación de Educación Superior.....	59
Ilustración 20 Opciones de Google Refine	64
Ilustración 21 Datos Inconsistentes sobre Organismos	64
Ilustración 22 Datos Inconsistentes sobre Normas.....	65
Ilustración 23 Forma Automática mediante Cluster	65
Ilustración 24 Datos Consistentes	66
Ilustración 25 Esquema de la aplicación	68
Ilustración 26 Código del Vocabulario LUDLAW	69
Ilustración 27 Servidor Apolo	70
Ilustración 28 EndPoint SPARQL.....	71
Ilustración 29 Resultado Normas	72
Ilustración 30 Resultado Normas sobre Derecho Educación	73
Ilustración 31 Resultado Países sobre derecho Autonomía Universitaria	74
Ilustración 32 Resultado Universidades del Sistema Educación	74
Ilustración 33 Resultado Normas que definen Organismos.....	75
Ilustración 34 Proceso de enlazado de datos (Gómez & Corcho, 2012)	76
Ilustración 35 Datasets sobre Países	77
Ilustración 36 Source para Enlazar	78
Ilustración 37 Dataset de DBpedia	79
Ilustración 38 Enlace de datos	79
Ilustración 39 DBpedia y Freebase	80
Ilustración 40 Inicializar el servicio Virtuoso	111
Ilustración 41 Pantalla inicial de Virtuoso	112
Ilustración 42 Login Virtuoso.....	112
Ilustración 43 Carga del RDF	113
Ilustración 44 Grafos RDF.....	113

Ilustración 45 Añadir Prefijos en Virtuoso	114
Ilustración 46 Resultado Tipo de Norma.....	115
Ilustración 47 Resultado Normas y Nombre del Presidente	115
Ilustración 48 Resultado Normas que Regularn Sistema Educación	116
Ilustración 49 Resultado Normas sobre Financiamiento Universitario	117
Ilustración 50 Resultado Organismos reguladores	117
Ilustración 51 Resultado Organismos de Consulta	118
Ilustración 52 Normas sobre Autonomía Universitaria.....	124

INDICE DE TABLAS

Tabla 1 Sentencia RDF.....	19
Tabla 2 Propiedades RDF.....	21
Tabla 3 Propiedades RDFS.....	22
Tabla 4 Clases y Propiedades AIIISO.....	26
Tabla 5 Resumen de características de las normas.....	48
Tabla 6: Metadatos de las normas	50
Tabla 7: Datos de los artículos	51
Tabla 8: Datos de organismos	52
Tabla 9 Glosario de Términos.....	56
Tabla 10 Vocabularios Reutilizados.....	61
Tabla 11 Búsqueda de organismos en Datasets.....	77
Tabla 12 Uris de los países	78
Tabla 13 Tripletas de las normas.....	123

RESUMEN EJECUTIVO

El presente proyecto de investigación se desarrolla con la finalidad de publicar datos enlazados en el ámbito de Legislación de Educación Superior en Latinoamérica que contiene información sobre: Universidades, Organismos Reguladores y Normas que regulan el Sistema de Educación Superior de cada país para cumplir con el objetivo del proyecto se utilizará los principios de Linked Data definidos por Tim Berners-Lee, de tal manera que la información puede ser accedida de forma pública.

Se construyó un vocabulario denominado LUDLAW el cual representa el ámbito del proyecto para ello se reutilizó vocabularios como: DUBLIN CORE, FOAF, DBPEDIA, BCNNORMS para contar con un vocabulario consensuado, además se utilizará tecnologías de la web semántica: RDF que permite contar con datos estructurados y que puedan ser entendidos por las máquinas.

Se desarrolló un programa en JENA que permite realizar la conversión a RDF; para realizar la publicación de datos se analizó diferentes repositorios de RDF como: Sesame, 4Store, AllegroGraph y Virtuoso. Mediante el ENDPOINT SPARQL que posee Virtuoso se realizó consultas con el fin de resolver las interrogantes planteadas al inicio de la investigación.

PALABRAS CLAVES: Linked Data, metadatos, RDF Schema, Web Semántica, SPARQL, URI, Jena, Recursos, Leyes, Educación Superior.

ABSTRACT

This research project is developed with the aim of publishing linked Legislation in the field of Higher Education in Latin America data that contains information about :University, Regulators and Regulations governing the higher education system of each country to meet the goal project Linked Data principles defined by Tim Berners- Lee , so that the information can be accessed is publicly used.

LUDLAW called a vocabulary which represents the scope of this project was reused as vocabularies built: DUBLIN CORE, FOAF , DBpedia , BCNNORMS to have a consensus vocabulary also semantic web technologies are used: RDF data that allows for structured and can be understood by machines.

Sesame, 4store, AllegroGraph and Virtuoso: for publishing RDF data from different repositories as discussed ; JENA a program that allows the conversion to RDF was developed. By having Virtuoso ENDPOINT SPARQL queries are performed in order to resolve the questions raised at the beginning of the investigation.

KEYWORDS: Linked Data, Metadata, RDF Schema, Semantic Web, SPARQL , URI , Jena, Resources , Law , Higher Education.

INTRODUCCIÓN

En años atrás la información era estática no existía interacción de los usuarios, es por ello que la web evoluciono a pasos agigantados y se pasó de la web 1.0 a la web 2.0 donde se podía interactuar.

En la web existe una gran cantidad de datos que no se encuentran enlazados, es decir estos datos son sueltos, lo que se desea obtener mediante el enlazado de datos es ofrecer búsquedas con mayor precisión, dotar de significado a los datos y mejorar la accesibilidad a los usuarios.

Actualmente diversas organizaciones publican conjuntos de datos enlazados, estos datos pueden ser accesibles desde cualquier parte del mundo mediante la web que ha evolucionado rápidamente.

Cuando se refiere a datos enlazados o Linked data, se menciona a datos publicados sobre la Web y de esta manera se busca cumplir con los siguientes principios fundamentales definidos por Tim Berners:

- Deben ser publicados sobre HTTP.
- Utilizar una URI para que identifique cada entidad existente.
- Brindar información sobre los recursos usando RDF.
- Incluir enlaces a otras URI esto permitirá el enlazado con otros datos existentes.(Bizer, Cyganiak, & Heath, 2007)

En la línea de datos gubernamentales se está haciendo un esfuerzo por publicar datos enlazados en distintos ámbitos por ejemplo salud, educación, legislación.(Cifuentes F. , 2011)

En este trabajo se propone publicar datos en el ámbito de Legislación de Educación Superior en Latinoamérica el cual abarca diferentes tipos de información: Normas de Educación Superior, Organismos de Educación, Universidades, Países, Sistema de Educación esta información se encuentra en dispersa en la web y en diferentes formatos como pdf, word y sitios web; para lograr el objetivo de publicación de datos se utilizará los principios de Linked Open Data diseñados por Tim Berners-Lee, de tal manera que estos puedan ser accedidos fácilmente y formen parte de la web de datos; esta

información puede ser estructurada y enlazada lo cual ayudará a que las búsquedas sobre Legislación Universitaria sea eficiente. También se utilizará diferentes tecnologías semánticas como: RDF, SPARQL, motores de almacenamiento de tripletas y herramientas para publicación de datos.

El desarrollo de la tesis comprende los siguientes capítulos:

El capítulo 1 abarca el Estado del Arte el cual permite conocer tecnologías, herramientas, estándares, principios e iniciativas de LinkedData que se pueden aplicar en el desarrollo del trabajo de investigación.

El capítulo 2 abarca el desarrollo de investigación sobre publicación de datos enlazados del ámbito sobre Legislación de Educación Superior, el mismo que se encuentra dividido en 6 fases:

- En la FASE 1 se detalla sobre la Identificación de fuentes y Extracción de Datos sobre Legislación de Educación Superior.
- En la FASE 2 se detalla el Desarrollo del Vocabulario sobre de Educación Superior y se reutiliza vocabularios existentes en este ámbito.
- En la FASE 3 se detalla sobre la Limpieza de Datos de Legislación de Educación Superior.
- En la FASE 4 se realiza la Generación de Datos a RDF mediante un programa desarrollado en Jena.
- En la FASE 5 se realiza la Publicación de Datos RDF en un RDF STORE.
- En la FASE 6 se identifica otras fuentes de datos con las cuales se puede Enlazar los datos sobre Legislación de Educación Superior.

Y al final del capítulo 2 contiene: discusión, conclusiones y recomendaciones que se pudieron obtener como resultado del desarrollo de la investigación.

OBJETIVOS

Objetivo General

Generar y publicar datos enlazados en el ámbito de Legislación de Educación Superior en Latinoamérica utilizando los principios de Linked Data.

Objetivos Específicos

- Extraer la información necesaria sobre Legislación de Educación Superior y almacenarla en una base de datos para luego convertirla en formato RDF.
- Analizar las tecnologías semánticas y proceso de publicación de datos enlazados relacionados con Linked Open Data y las herramientas para la publicación de datos RDF para aplicar en el proyecto de investigación.
- Reutilizar vocabularios existentes para la construcción del vocabulario sobre Legislación de Educación Superior.
- Utilizar tecnologías y herramientas que permita convertir los datos a formato RDF.
- Realizar consultas SPARQL sobre los datos de Legislación de Educación Superior con la finalidad de resolver las interrogantes planteadas al inicio de la investigación.

PROBLEMA

En la actualidad se están realizando varios proyectos sobre web semántica en diferentes áreas de conocimiento algunas como: educación, bioinformática, salud y legislación(Cifuentes F. , 2011); basado en este impulso se ha iniciado el proyecto de publicación de datos enlazados en el ámbito de Legislación de Educación Superior en Latinoamérica que comprende:

- Normas (Ley, Reglamento, Constitución, Decreto, Resolución, Ordenanzas) sobre la Educación Superior en diferentes países de Latinoamérica.
- Organismo de Educación Superior encargados de regular y acreditar la calidad de la educación en cada país.
- Universidades que forman parte del sistema de Educación Superior.

Esta información actualmente se encuentra en formato no estructurado debido a la falta de acuerdos en un esquema de representación lo cual dificulta el procesamiento de información de forma automática.

Para solucionar el problema de formatos de la información se pretende convertirla a formato estructurado denominado RDF que es el lenguaje de la web semántica, mediante la utilización de tecnologías semánticas se busca dar solución.

1. ESTADO DEL ARTE

1.1. Introducción.

Internet ha pasado de un espacio informativo conectando textos, hacia una red proveedora de conexiones sociales, luego a un gran repositorio de conocimiento y se dirige a un espacio que conectará entidades virtuales con lugares y cosas reales; con el pasar de los años la Web ha evolucionado a pasos agigantados desde la Web 1.0, Web 2.0 y ahora en la actualidad la Web de Datos Enlazados la cual trata de vincular los distintos datos que están distribuidos en la Web con el propósito de mejorar la búsqueda de información.

La nueva era es la Web Semántica que es una Web extendida, dotada de mayor significado en la que cualquier usuario en Internet podrá encontrar respuestas a sus preguntas de forma más rápida y sencilla gracias a que la información está mejor definida.

Según (Bizer, Cyganiak, & Heath, 2007), supone una evolución de la Web hacia un espacio global de información en el que la navegación se realiza a través de datos estructurados enlazados en vez de realizarse a través de documentos, como sucede ahora. Así se pasa de una Web basada en documentos HTML, en la que el usuario es el destinatario de la información publicada, a una Web de Datos Enlazados que están expresados en RDF.

Mediante la utilización de Linked Data que es un conjunto de principios para exponer datos enlazados en la Web y no solamente páginas web que contienen información, se logra tener datos enlazables, interoperables, visibles y reusables con el fin de dar un valor agregado a la web actual.

1.2. Evolución de la Web.

La historia de la Web(Rubio, 2006)se remonta hace más de 40 años. La Corporación RAND comenzó la investigación financiada por la Fuerza Aérea de los EE.UU. para determinar cómo desarrollar redes de gran escala, redes potentes, comunicación distribuida para el comando y control militar.

La evolución de la World Wide Web(Berners-Lee, 2007), se ha producido en una serie de elementos de muy diverso signo.

Desde el diseño en las interfaces gráficas de las propias páginas web y de los navegadores, pasando por las funcionalidades de la Web actual que permite:

- Transacciones financieras y de e-comercio
- Múltiples recursos multimedia
- La potencia de los buscadores
- Bases de datos
- Los lenguajes hipertextuales que permiten numerosas aplicaciones por medio de la información estructurada, hasta desembocar en la explosión de la Web.

En la primera generación de la Web el objetivo fundamental consistía en buscar recursos basándose en patrones, materia que ha sido de amplio estudio por la disciplina de recuperación de la información(Oreilly, 2012).

En cambio en la segunda generación de la Web o también llamada Web2.0 o Web Social, se toma en cuenta para la búsqueda inicial de recursos algunas opciones como:

- Colaborar en los recursos esto quiere decir utilizar ambientes de edición distribuida. Ejemplo: Wikipedia¹.
- Compartir los recursos. Ejemplos: Blogs
- Clasificar los recursos, mediante el uso de etiquetas o tags. Existen múltiples servicios que permiten esta opción y depende del tipo derecurso que se está clasificando, por nombrar algunos: delicious², flickr³.
- Comentar los recursos, añadiendo hilos de discusión.
- Calificar los recursos, agrupando en una lista de favoritos.

La web hoy en día se ha convertido en un medio flexible para la comunicación, la educación, comercio, entretenimiento, información entre otros servicios que brinda esta tecnología(Berners-Lee, 2012).

La evolución de la web y tecnologías se ha dado a pasos agigantados:

- Las primeras tecnologías básicas: HTML⁴ y HTTP⁵, hasta hoy en día.
- Y las nuevas tecnologías que aparecieron como: Java⁶, JavaScript⁷, ASP⁸, JSP⁹, PHP¹⁰, XML¹¹, algunas de las más conocidas, que

¹<http://es.wikipedia.org>

²<http://delicious.com/>

³<http://www.flickr.com/>

⁴<http://www.w3.org/MarkUp>

⁵<http://www.w3.org/Protocols>

permiten una web mejor, más amplia, más potente, más flexible, o más fácil de mantener.

En la ilustración 1 observamos la evolución de la web:

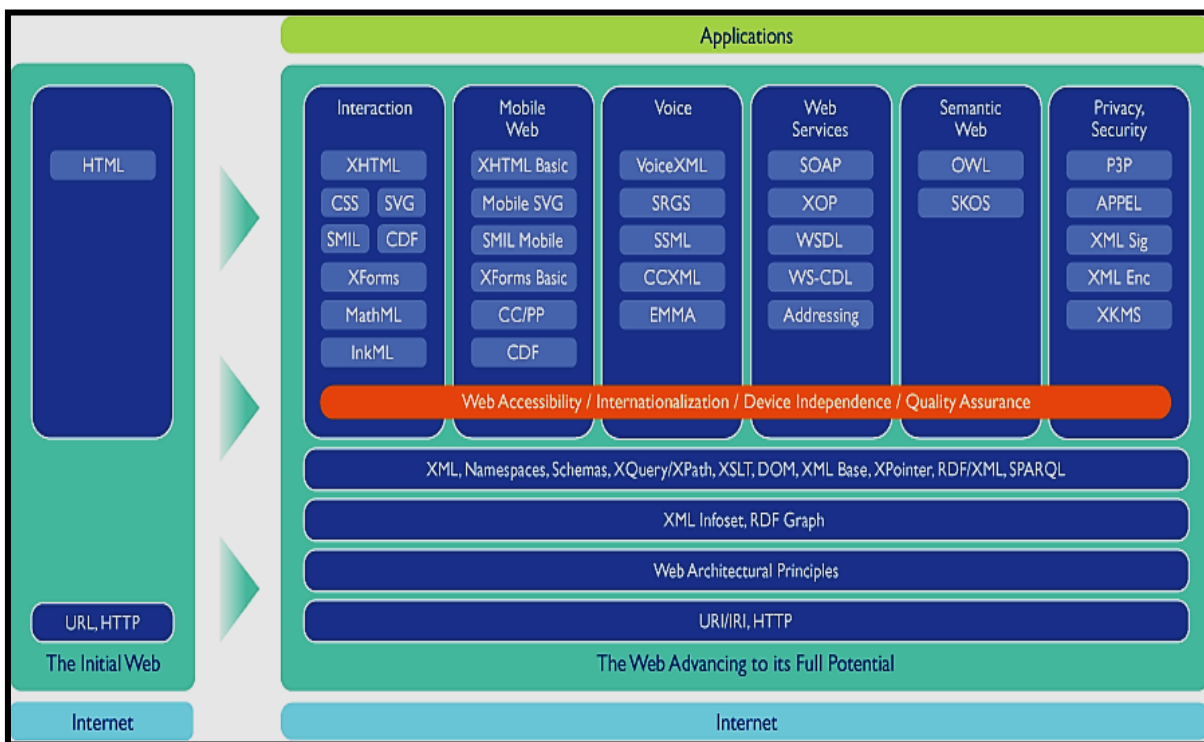


Ilustración 1 Evolución de la Web (W3C12)

1.3. Web Semántica.

Los orígenes de la “Web semántica”, por su parte, se remontan al año 2001, cuando el científico británico Tim Berners Lee, Inventor de la World Wide Web y el investigador en inteligencia artificial James Hendler del Departamento de Ciencias de la Computación de la Universidad de Maryland, presentaron un artículo en la revista Nature abordando por primera vez la posibilidad de generar este tipo de Web (Tim Berners & Hendler, 2001).

La Web debería estar diseñada no sólo como espacio de información accesible para los seres humanos sino también como espacio donde pudieran acceder

⁶ http://www.java.com/es/download/faq/whatis_java.xml

⁷ https://developer.mozilla.org/en/About_JavaScript

⁸ <http://www.asp.net/get-started>

⁹ http://es.wikipedia.org/wiki/JavaServer_Pages

¹⁰ <http://www.php.net/>

¹¹ <http://www.w3.org/XML/>

¹² <http://www.w3c.es/Presentaciones/2005/1018-WebSemanticaREBIUN-MA/img/WebEvolution.png>

las computadoras. El objetivo del acceso por parte de las computadoras sería ayudar a los seres humanos a buscar y conseguir la información que necesitan, con la mayor rapidez y garantía de éxito posibles.

Según Berners-Leela web semántica(Berners-Lee T. , 2012)propone superar las restricciones o limitaciones de la web actual mediante la introducción de descripciones explícitas del significado, la estructura interna y la estructura global de los contenidos y servicios disponibles en la WWW.

El objetivo fundamental de la web semántica es crear un medio universal que permita el intercambio de datos y brindar mayor significado a la misma para que puedan ser interpretadas por las máquinas.

1.3.1. Concepto de Web Semántica.

La Web semántica según el W3C: “La Web semántica proporciona un marco común que permite que los datos sean compartidos y reutilizados a través de aplicaciones, empresas y fronteras comunitarias. Está basado en Resource Description Framework (RDF) e integra una variedad de aplicaciones utilizando XML para la sintaxis y URI para las denominaciones”. Es una Web extendida, dotada de mayor significado en la que cualquier usuario en Internet podrá encontrar respuestas a sus preguntas de forma más rápida y sencilla gracias a una información mejor definida.

Para entender un poco más sobre Web Semántica, se presenta un ejemplo de un buscador actual como en la ilustración 2, Google que es capaz de mostrar varios resultados. Para este ejemplo se busca información sobre el Papa Juan Pablo II, frente a un buscador semántico como WolframAlpha que arroja resultados que son sin duda efectivos y acoplados a la necesidad del usuario se puede observar en la Ilustración 3.

En el buscador de Google nos arrojó cerca de 14.2000.000 resultados sobre esta búsqueda un tiempo muy corto de 0,36 segundos. Si analizamos un poco los resultados nos damos cuenta que tendremos que navegar por ellos para elegir el adecuado.



Ilustración 2 Buscador Actual

En la ilustración 3 podemos observar que los resultados que presenta el buscador WolframAlpha desarrollado por la compañía Wolfram Research utiliza la tecnología Question Answer para la recuperación de la información. El buscador responde directamente a las preguntas de los usuarios y no presenta una lista de resultados. Presenta un resumen sobre la información de Juan Pablo II.

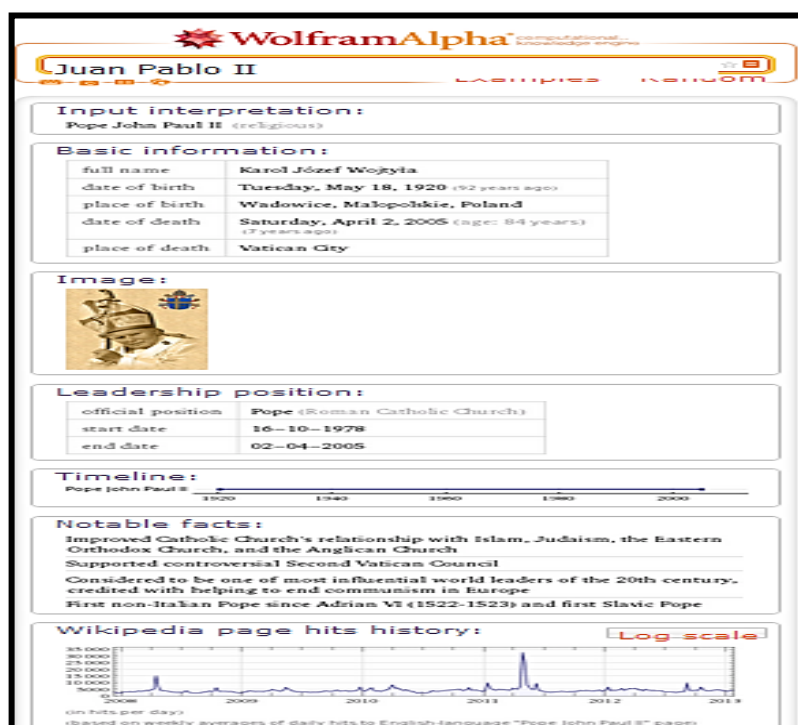


Ilustración 3 Buscador Semántico

1.3.2. Campos de aplicación de la Web Semántica.

Según(Castells & Camacho, 2006), existen diferentes campos de aplicación donde las ideas de la web semántica pueden tener utilidad:

- Gestión del conocimiento
- Comercio electrónico
- Búsqueda de información en la web
- Enseñanza
- Librerías digitales
- Turismo y Patrimonio cultural

En la ilustración 4 se presenta un diagrama en capas según Tim Berners-Lee, en si es la infraestructura de tecnologías y lenguajes necesarios para la implementación de la Web Semántica:

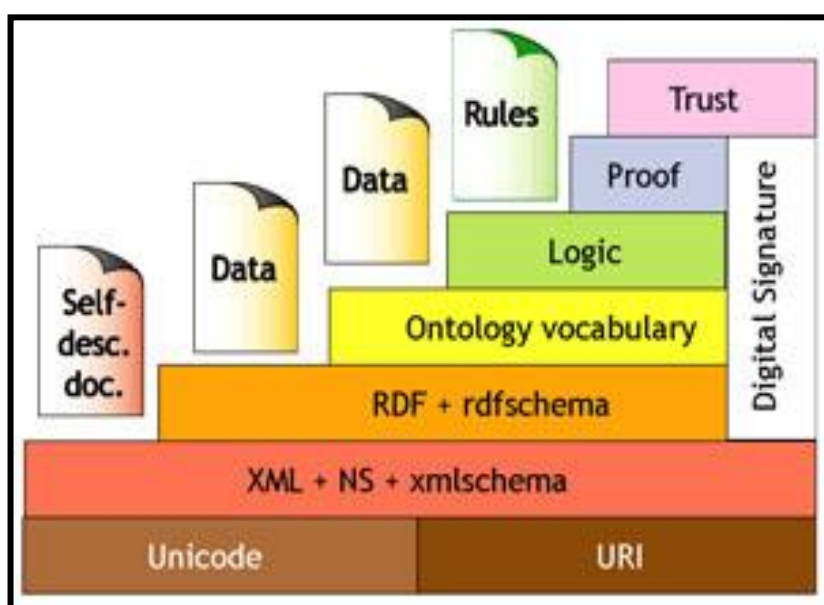


Ilustración 4 Arquitectura de la Web Semántica (W3C¹³)

¹³<http://www.w3.org/2000/Talks/1206-xml2k-tbl/sweb-stack.gif>

1.3.3. Descripción de cada uno de las capas de la web semántica.

Capa 1

- **Unicode:** es un sistema internacional estándar que proporciona un número único para cada carácter, sin importar la plataforma ni el programa.
- **URI:** es un sistema de direccionamiento e identificación de recursos.

Capa 2

- **XML:** es un sistema que permite definir lenguajes de marcas para usos específicos.
- **NS:** permite combinar diversos lenguajes de marcado creados con XML en un mismo documento.
- **XML SCHEMA:** sirve para definir tipos de documentos complejos en los que se pueden especificar tipos de datos, listas de componentes y restricciones.

Capa 3

- **RDF + RDFSCHEMA:** define un lenguaje universal con el que se puede expresar diferentes ideas en la Web Semántica.

Capa 4

- **Ontológica:** sirve para clasificar la información agregando clases y propiedades a los recursos.

Capa 5

- **Lógica:** se refiere al estudio de las reglas formales para las ontologías.

Capa 6

- **Proof:** Se intercambiarán “pruebas” escritas en el lenguaje unificador de la Web Semántica.

Capa 7

- **Trust:** debe servir para otorgar confianza a las transacciones en la Web a través que se llevarán a cabo no solamente entre usuarios y sitios web sino también entre programas de software
- **Digital Signature:** proporcionará soporte específico a esta capa, tal como muestra el diagrama.

1.3.4. Ventajas y desventajas de la Web Semántica.

La web semántica brinda las siguientes ventajas(Salazar, 2011):

- Permite organizar la gran cantidad de información que se encuentra suelta y redundante en la Web.
- Ayuda a reducir costos y tiempo en localizar información útil en la Web, ya que actualmente se debe realizar el análisis semántico de la información.
- Incorpora contenido semántico a las páginas que se suben a Internet. Esto permite una mejor organización de la información, asegurando búsquedas más precisas por significado y no por contenido textual.

La web semántica así como posee ventajas también existen algunas desventajas:

- Es costoso adaptar los documentos de Internet, para que puedan ser procesados al lenguaje de la Web Semántica.
- Sobrecarga de información.
- Heterogeneidad de las fuentes de información.

1.4. Linked Data.

Linked Data (Bizer, How to publish linked data on the web, 2012) permite construir la Web de los datos. Es una gran base de datos interconectados y distribuidos en la Web. Los datos se enlazan y se examinan de una forma similar a la utilizada para vincular los documentos de tipo HTML.

El término "datos enlazados" hace referencia al método con el que se pueden mostrar, intercambiar y conectar datos a través de URIs en la Web.

Un ejemplo de un gran conjunto de datos es DBPedia lo que contiene es información de Wikipedia que está en formato disponible en RDF.

Además DBPedia no sólo incluye datos de Wikipedia, sino también incorpora enlaces a otros dataset o conjuntos de datos en la Web, por ejemplo, a Geonames(W3C, 2012).

La ilustración 5 se muestra los diferentes datasets o conjuntos de datos de diversos tipos organizados por los dominios, estos datos se encuentran conectados entre sí de forma que componen la nube de Linked Data.

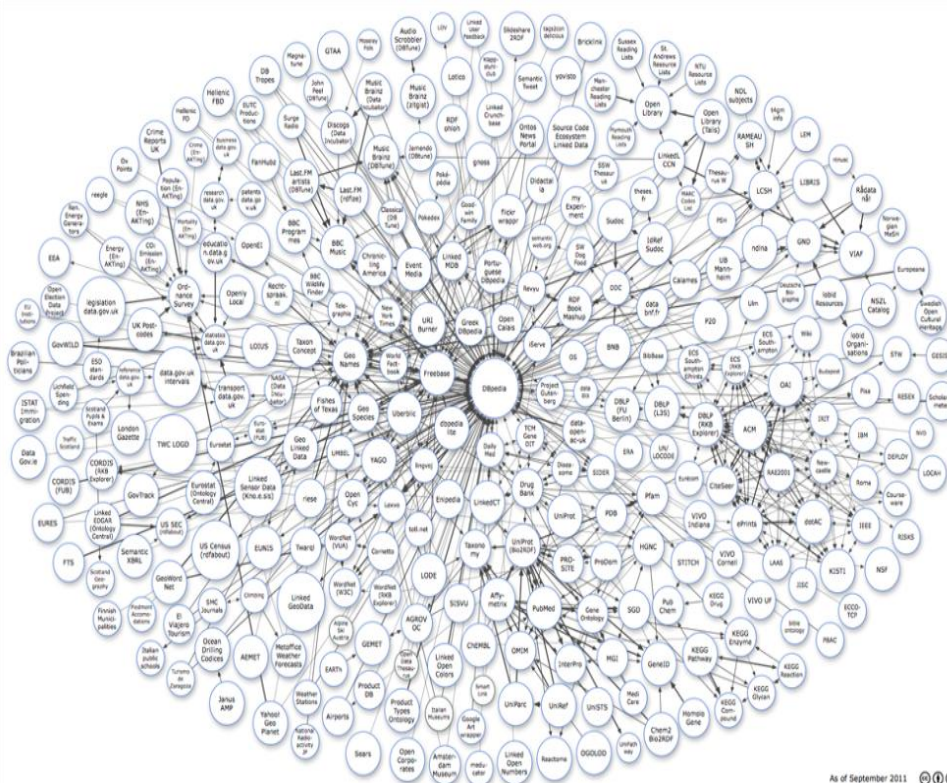


Ilustración 5 Nube de Linked Data (W3C, 2011)

1.4.1. Principios de Linked data.

Linked Data se basa en la aplicación de cuatro principios básicos que fomentarán el crecimiento de la Web, tanto a nivel de los documentos HTML, como a nivel de los datos expresados en RDF o Web Semántica. Según Berners-Lee se definió cuatro principios para la publicación de Linked data (Bizer, Heath, & Berners, 2009):

1. Usar URIs (Uniform Resource Identifiers) identificando los recursos de forma unívoca;
2. Usar URIs HTTP con este uso se pretende asegurar que cualquier recurso pueda ser buscado y accedido en la Web.
3. Ofrecer información sobre los recursos usando RDF, se debe obtener información útil sobre dicho recurso, representada mediante descripciones estándares en RDF.
4. Incluir enlaces a otros URIs, facilitando el vínculo entredistintos datos distribuidos en la web. Mediante los enlaces a recursos provenientes de sitios más especializados en determinados dominios se generara valor añadido a la información que se brinda.

Con la utilización de estos principios se podrá obtener datos interconectados y así ofrecer valor agregado a la Web.

1.4.2. Proceso para publicación de datos enlazados.

Según (Corcho & Gómez, 2010) menciona que el proceso publicación de datos enlazados en la Web incluye los siguientes pasos que se describen a continuación:



Ilustración 6 Proceso de publicación de datos enlazados(Corcho & Gómez, 2010)

- 1. Identificar y seleccionar las fuentes de datos.-** los datos pueden ser provenientes de diferentes sitios como: base de datos, documentos web, hojas de Excel, documentos de Word.
- 2. Desarrollo de vocabularios.-** para la construcción es recomendable reutilizar los vocabularios existentes con el fin de crear uno que satisfaga las necesidades propuestas. Entre los más reconocidos y utilizados encontramos a FOAF, Dublin Core, Skos, Aiiiso, Bio y muchos más dependiendo del dominio que se desee representar y accedidos mediante servicios de consulta.
- 3. Limpieza de datos.-** es necesario antes de realizar la conversión y publicación de los datos testear y depurar con el fin que la información sea correcta.

4. **Generación de datos RDF.**-una vez identificada la información necesaria para responder a las preguntas planteadas, se debe convertir la información en un lenguaje entendible por maquinas llamado RDF.

5. **Publicación de datos RDF.**- la publicación de la información en formato RDF o tripletas es aconsejable almacenar en un RDF Store.

6. **Enlazar datos RDF con otras fuentes.**-luego de haber realizado la publicación de los datos debemos enlazar los datos con otros ya existentes en la web de datos y de esta manera se genera un valor agregado.

1.5. Tecnologías Semánticas.

1.5.1. URI.

Es un identificador uniforme de recurso, es definido por el estándar RFC 3986 (Berners-Lee T. , 2005) y se define como una cadena de caracteres que identifica unívocamente a un recurso sobre un sistema.

Los componentes de una URI según el estándar (RFC3986)¹⁴:

ESQUEMA: PARTE JERARQUICA ? SOLICITUD # FRAGMENTO

- **Esquema:** El esquema identifica el protocolo que se va a usar. El más utilizado es HTTP.
- **Parte jerárquica:** contiene la información del dominio o IP para acceder al servidor y la ruta en el servidor para acceder al recurso.
- **Solicitud:** la solicitud indica variables que se pasan al recurso web.
- **Fragmento:** permite indicar una subdirección dentro del recurso al que apunta la dirección.

Ejemplo:

http : //www.hispamedios.com/ ? pagina=2 # inicio

Una URI se diferencia de una URL (Uniform Resource Locator) porque la primera permite identificar recursos mediante fragmentos dentro de un documento mientras que la segunda no.

1.5.2. RDF.

RDF es el marco de descripción de recursos para metadatos en la Web elaborado por el W3C (Méndez, 2012). Se basa en la idea de declarar recursos usando la expresión en la forma sujeto-predicado-objeto. Esta expresión es conocida en la terminología RDF como triple o triplete. RDF se basa en la idea de identificar recursos utilizando identificadores Web llamados URIS-Identificadores Uniformes de Recursos, y de describir a los mismos en términos de propiedades sencillas y valores de propiedades.

¹⁴<http://www.ietf.org/rfc/rfc3986.txt>

1.5.2.1. Elementos primordiales de RDF.

RDF se basa en 3 elementos primordiales(W3C, 2004):

- **Recursos:** son todas las cosas descritas por expresiones RDF, desde una página web completa como un documento HTML o una colección completa de páginas, hasta una parte de una página web.
- **Propiedades:** una propiedad es un aspecto específico, característica, atributo o relación que puede utilizarse para describir un recurso.
- **Declaraciones o sentencias:** una declaración RDF es una propiedad más el valor de dicha propiedad para un recurso específico.

1.5.2.2. Sentencia RDF.

Una sentencia RDF está compuesta por 3 partes(Méndez, 2012):

- **Sujeto:** una referencia URI, una persona, o un nodo, es el ente al cual nos referimos
- **Predicado:** es la propiedad o relación que se desea establecer acerca del sujeto
- **Objeto:** valor de la propiedad, o un recurso que puede considerarse el valor definido.

Por ejemplo:

Ora Lassila es el creador [autor] del recurso: <http://www.w3.org/Home/Lassila>.

Esta sentencia comprende las siguientes partes:

Tabla 1 Sentencia RDF

Sujeto (Recurso)	http://www.w3.org/Home/Lassila
Predicado (Propiedad)	Creator
Objeto (literal)	"Ora Lassila"

Los grafos RDF son conjuntos de tripletas:



Ilustración 7 Grafo RDF

Elaboración: María Fernanda Valverde

Ejemplo de RDF:

```
<rdf:RDF
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:dc="http://purl.org/dc/elements/1.1/">
<rdf:Descriptionrdf:about="http://www.w3.org/Home/Lassila">
<dc:creator>Ora Lassila</dc:creator>
</rdf:Description>
</rdf:RDF>
```

1.5.2.3. Áreas de aplicación de RDF.

Algunas de las áreas de aplicación de RDF se muestran a continuación:

- Búsquedade recursos: proporcionar mejores capacidades de los motores de búsqueda.
- Catalogación: describir el contenido y las relaciones de contenido disponibles en un sitio Web.
- Agentes de software inteligentes: para facilitar el intercambio de conocimientos
- Para describir los derechos de propiedad intelectual de las páginas Web
- Para expresar las preferencias de privacidad de un usuario
- Firmas Digitales

1.5.2.4. Propiedades RDF.

Describen el concepto de una propiedad RDF como una relación entre los recursos en cuestión y los recursos de los objetos.

Tabla 2 Propiedades RDF

NOMBRE DE LA PROPIEDAD	COMENTARIO
rdf: type	Se utiliza para clasificar recursos en categorías o clases.
rdf:Property	Se utiliza para clasificar todas las propiedades de cada una de las clases.

1.5.3. RDFS.

Permite definir los términos que se usarán en las declaraciones RDF y otorgara significados específicos. Para evitar definiciones conflictivas del mismo término, RDF utiliza los namespaces de XML.

RDFS(W3C, 2008)permite modelar metadatos con una representación explícita de su semántica y permite especificar restricciones de tipos de datos para los sujetos y objetos de las tripletas de RDF, introduciendo unas primitivas de modelado orientado a objetos como por ejemplo: rdfs:subclassOf,rdfs:subPropertyOf.Un schema define el significado, características y relaciones de un conjunto de propiedades.

En la tabla 3 podemos observar las propiedades RDFS clasificadas por algunos criterios:

Tabla 3 Propiedades RDFS

CLASIFICACION	NOMBRE DE CLASE	COMENTARIO
Básicas	rdfs: Resource	Es la clase a la que pertenecen todos los recursos.
	rdfs: Class	Permite declarar recursos como clases para otros recursos.
	rdfs:Literal	Es la clase que contiene valores literales, cadenas y enteros.
	rdfs:Datatype	Esta clase abarca los tipos de datos para RDF.
Define relaciones	rdfs:subclassof	Se aplica a las instancias que tienen una propiedad valorada por la clase.
	rdfs:subPropertyOf	Se aplica a las propiedades que pueden ser interpretadas como un subconjunto de otras propiedades, permitiendo la inferencia entre conceptos.
Restricciones de propiedades	rdfs: Domain	
	rdfs: Range	Se aplican a propiedades que son valorados por clases. Una propiedad no puede tener más de un valor para rdfs:range.
Documentación	rdfs: Comment	Se usa para proporcionar una descripción de un recurso.
	rdfs:label	Se usa para proporcionar una versión claramente entendible del nombre de un recurso.

1.5.4. SPARQL.

Es lenguaje de consultas para RDF y en la actualidad es recomendado por W3C. Mediante su expresividad permite realizar consultas a múltiples fuentes de datos o grafos, que deben estar en RDF. Su sintaxis es similar a la del lenguaje SQL aunque orientado a tripletas y grafos RDF. (DuCharme, 2011)

Los resultados de las consultas SPARQL pueden ser conjuntos de tripletas RDF.

1.5.4.1. Sintaxis SPARQL.

SPARQL mantiene una sintaxis similar a SQL, mantiene una estructura declarativa con la típica sintaxis basada en cláusulas(Classora Technologies, 2012).

- PREFIX.- es equivalente al namespace en XML, en lugar de usar un url varias veces se utiliza un prefix.
- SELECT.- indica cada elemento a seleccionar: sujeto, predicado, objeto. Precedido de un símbolo ?. De igual forma como en SQL se puede utilizar * para representar todos los elementos.
- FROM.- la tabla a consultar (en este caso, la ontología o grafo a consultar) no se especifica mediante la cláusula FROM, sino que se indica de manera externa a la consulta.
- WHERE.- siendo la parte más importante de la consulta. Para filtrar triplete, es posible establecer una parte de la tripleta (sujeto, predicado, objeto) como variable, asignándole un nombre precedido del símbolo ?.

Ejemplo de consulta SPARQL:

```
SELECT *  
WHERE  
{ ?s a ?o .filter (regex(?o,"^http://freebase.*country")) } limit 10
```

1.6.Vocabularios.

Con la finalidad de que sea posible procesar los datos es necesario utilizar los términos de conocidos vocabularios siempre que sea posible y si no se debe definir los términos requeridos en un nuevo vocabulario que abarque toda la información que se desea representar. Existen algunos vocabularios los cuales pueden ser reutilizados dependiendo de la necesidad para procesar los datos:

1.6.1. DUBLIN CORE.

Es un modelo de metadatos elaborado por DCMI (Dublin Core Metadata Initiative), una organización dedicada a fomentar la adopción extensa de los estándares interoperables de los metadatos y a promover el desarrollo de los vocabularios especializados de metadatos para describir recursos para permitir sistemas más inteligentes del descubrimiento del recurso. Las implementaciones de Dublin Core (DCMI, 2012), usan generalmente XML y se basan en el Resource Description Framework.

Dublin Core define 15 definiciones semánticas descriptivas que son necesarias para describir, identificar, procesar, encontrar y recuperar un recurso en la web.

Algunas de las propiedades con las que cuenta Dublin Core son: title, subject, description, language, creator, Publisher, date.

Cualquier persona puede utilizar los metadatos de Dublin Core para describir los recursos de un sistema de información. Actualmente las páginas Web son uno de los tipos más comunes de recursos que utilizan las descripciones de Dublin Core. Algunos grupos de interés que utilizan los metadatos de Dublin Core son: bibliotecas, instituciones del gobierno, organizaciones educativas, sector científico de la investigación y autores de páginas.

Ejemplo:

```
<rdf:RDF
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:dc="http://purl.org/dc/elements/1.1/">
<rdf:Description rdf:about=" http://dublincore.org/documents/dces/">
<dc:creator> Dublin Core Metadata Initiative</dc:creator>
<dc:title> Dublin Core Metadata Element Set, Version 1.1</dc:title>
<dc:description> This document provides ready reference for the Dublin Core Metadata
Element Set, Version 1.1. For more detailed documentation and links to historical
versioning information, see the document "DCMI Metadata Terms".</dc
<dc:date> 2012-06-14</dc:date>
</rdf:Description>
</rdf:RDF>
```

1.6.2. FOAF (Friend of a Friend)¹⁵.

FOAF es un vocabulario/ontología para describir organizaciones, personas sus actividades y expresar sus relaciones con otras personas y objetos. Este vocabulario permite trabajar sobre información personal de forma sencilla y simplificada con la finalidad de ser procesada, reutilizada y compartida. El vocabulario FOAF está compuesto por clases y propiedades, algunas de ellas son:

Clases:

- Person, Organization
- Group, Document
- Image

Propiedades:

- Name, mbox
- Nick, gender

La URI de este vocabulario es: <http://xmlns.com/foaf/spec/>

Ejemplo:

```
<foaf:Person rdf:about="#danbri" xmlns:foaf="http://xmlns.com/foaf/0.1/">
<foaf:name>Dan Brickley</foaf:name>
<foaf:homepage rdf:resource="http://danbri.org/" />
<foaf:openid rdf:resource="http://danbri.org/" />
<foaf:imgrdf:resource="/images/me.jpg" />
</foaf:Person>
```

Las palabras que están en negrita son algunos de los términos que se encuentra en el vocabulario FOAF.

¹⁵<http://xmlns.com/foaf/spec/index.rdf>

1.6.3. AIISO(Academic Institution Internal Structure Ontology).

AIISO¹⁶ es un vocabulario para describir la estructura organizacional interna de una institución académica. Este vocabulario está diseñado para trabajar en relación con otros vocabularios como FOAF, AIISO-ROLES se utiliza para describir el rol que desempeñan las personas dentro de la institución(Styles & Shabir, 2008).

La URI de este vocabulario es:<http://purl.org/vocab/aiiso/schema#>

El vocabulario AIISO se encuentra definido por 15 clases y 10 propiedades desde las cuales se presenta un extracto en la tabla 4.

Tabla 4 Clases y Propiedades AIISO

NOMBRE	TIPO	DEFINICIÓN
Center	Class	Es un grupo de personas reconocidas por una organización.
College	Class	Es un grupo de personas reconocidas por una organización como un grupo formal, que se refiere por la organización en un collage.
Course	Class	Grupo de conocimiento que representa una colección de materiales educativos
Department	Class	Grupo de personas reconocidas por una organización como un grupo formal, referidas por la organización como un departamento.
code	Property	Se utiliza un código para referirse a una institución.
description	Property	Breve descripción de un grupo de conocimiento u organización.
name	Property	El nombre dado a un grupo de conocimiento u organización.
teaches	Property	Grupo de personas que se dedican a enseñar.
part_of	property	Parte de una organización.

¹⁶<http://vocab.org/aiiso/schema>

1.7.Repositorios de datos RDF.

Actualmente existen diferentes repositorios para almacenar las tripletas conocidas con el nombre de RDF Store o Triple Store. A continuación se mencionan algunas de las más utilizadas:

1.7.1. AllegroGraph.

AllegroGraph es un sistema moderno y de alto rendimiento, base de datos gráfica persistente. AllegroGraph utiliza la utilización eficiente de la memoria en combinación con el almacenamiento basado en discos, lo que le permite escalar hasta miles de millones de quads, mientras que el mantenimiento de un rendimiento superior. AllegroGraph apoya SPARQL, RDFS + razonamiento +, y el prólogo de numerosas aplicaciones cliente(Franz Incorporated, 2011).

Características:

- Soporta SPARQL
- Recuperabilidad completo y rápido para las consultas, se puede manipular los datos en triple-Store por diferentes interfaces y lenguajes como JAVA, HTTP y LISP.
- Indexación dinámica y automática
- Soporta RDFS++
- Carga de archivos en el formato Turtle.

1.7.2. Virtuoso.

Virtuoso es una herramienta para almacenar tripletas, las principales características son(OpenLink Software, 2010):

- Gestión de datos RDF
- Servidor de Linked Data
- Despliegue de servicios REST o SOAP

Características:

- Tiene licencia GLP y también posee licencia pagada para otras versiones del producto.
- Plataforma: Windows, Linux, Mac OS X

- Posee un diccionario de datos en donde se almacena toda la información de los objetos de los usuarios.
- Ofrece un endpoint SPARQL que permite la consulta de recursos contenidos en el servidor.

1.7.3. 4store¹⁷.

Es una base de datos RDF eficiente, escalable y estable, fue diseñado por Steve Harris y desarrollado en Garlik para respaldar sus aplicaciones de Web Semántica. Se ha estado proporcionando la plataforma de base de alrededor de 3 años.

Características:

- Soporte para datos RDF
- Escrita en C
- Disponible bajo GPL
- Ofrece un endpoint SPARQL que permite la consulta de recursos contenidos en el servidor.
- Posee escalabilidad y estabilidad sobre los datos.
- Plataforma: Unix, Mac OS X.

1.7.4. Sesame¹⁸.

Es un marco extensible de Java que soporta inferencia y búsqueda sobre RDF Schema. Se puede implementar como un servidor web o utilizarse como una biblioteca de Java. Las características incluyen varios idiomas de la consulta (SeRQL y SPARQL), el apoyo a la inferencia y memoria RAM, disco o de almacenamiento de RDBMS. Además, la API central de Sesame es el almacenamiento independiente y son apoyados por muchos otros fabricantes de bases de datos RDF (Sesame, 2010).

¹⁷<http://4store.org/>

¹⁸<http://sourceforge.net/projects/sesame/>

Características

- Soporte completo para SPARQL 1.1 Consulta y actualización.
- Streaming analizadores RDF y Escritores
- Flexible independiente del proveedor de la API para el acceso a muchos tipos de bases de datos RDF
- Ligerero RDF interfaz gráfico
- Configuración modular accesible a través de Maven

1.8. Herramienta para generación de RDF: JENA.

Jena¹⁹ es un API para Java, es decir, un entorno para el desarrollo de aplicaciones en el lenguaje de programación Java para la web semántica.

Es un Framework desarrollado por HB Labs para manipular metadatos desde una aplicación Java. En la actualidad existen dos versiones(Jena, 2011):

Jena 1:

- Principalmente soporte para RDF.
- Capacidades de razonamiento limitadas.

Jena 2:

- Incluye además una API para el manejo de ontologías.
- Soporta el lenguaje OWL.

1.8.1. Características.

Jena permite gestionar todo tipo de ontologías (añadir hechos, borrarlos y editarlos), almacenarlas y realizar consultas contra ellas. Soporta RDF, DAML y OWL y es independiente del lenguaje. Los recursos no están ligados estáticamente a una clase java particular.

Incluye varios componentes:

- ARP: un parser de RDF
- API RDF
- API de Ontologías para OWL, DAML y RDF Schema.
- Subsistema de razonamiento
- Soporte para persistencia

¹⁹<http://jena.apache.org/>

- RDQL: Lenguaje de consultas de RDF

API RDF de Jena

Permite crear y manipular modelos RDF desde una aplicación Java.

Proporciona clases java para representar:

- Modelos: son conjuntos de statements.
 - Creación de modelos
 - Escritura y lectura de modelos
 - Carga en memoria de modelos
 - Navegar un modelo a partir de la URI de un recurso
 - Consultar un modelo: se podrá buscar información del modelo y realizar consultas avanzadas.
 - Operaciones sobre modelos: unión, intersección y diferencia.
- Recursos, Propiedades, Literales y Statements.

1.9. Metodologías para la construcción de ontologías.

Existen diferentes propuestas metodológicas para la construcción y diseño de ontologías, entre las más utilizadas están: Methontology de la Universidad Politécnica de Madrid y Neón de un proyecto europeo (Barrera & Núñez, 2012).

1.9.1. Metodología Methontology.

Esta metodología permite la construcción de ontologías a nivel de conocimiento incluyendo la identificación del proceso de desarrollo de la ontología, un ciclo de vida basado en la evolución de prototipos y técnicas particulares para realizar cada actividad (Fernandez, Gomez, & Juristo, 1997).

El ciclo de vida de la metodología se lo puede observar en la ilustración 9, donde las actividades de gestión y de soporte se realizan simultáneamente con las actividades de desarrollo.

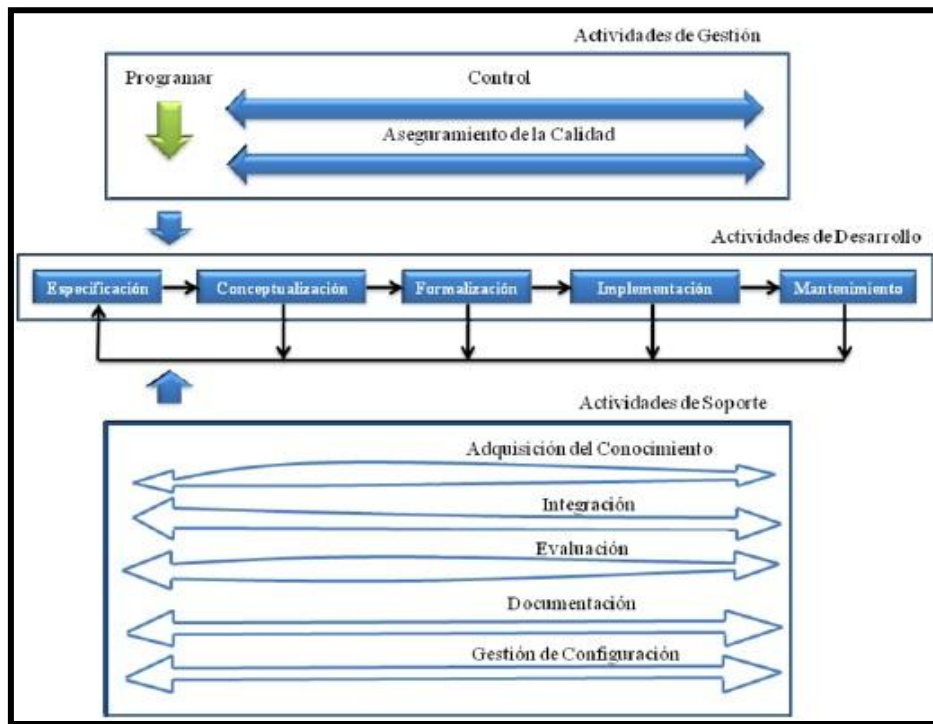


Ilustración 8 Ciclo de vida Methontology (Fernandez, Gomez, & Juristo, 1997)

1.9.2. Metodología Neón.

Neon²⁰ es una metodología ontológica para la construcción de redes de ontologías basada en escenarios que se apoya en los aspectos de colaboración de desarrollo de ontologías y la reutilización, así como en la evolución dinámica de las redes de ontologías en entornos distribuidos. El objetivo es mejorar la capacidad de manipular múltiples ontologías en red, incluyendo evolución y mantenimiento de las mismas, además de orientar el proceso de desarrollo de ontologías colaborativas y la reutilización de recursos ontológicos y no ontológicos (Neon, 2010). Los escenarios que pueden surgir durante el desarrollo de ontologías y redes de ontologías son los siguientes:

- Escenario 1. Desarrollo de redes de ontologías desde la especificación hasta la implementación.
- Escenario 2. Desarrollo de redes de ontologías mediante reutilización y reingeniería de recursos no ontológicos.
- Escenario 3. Desarrollo de redes de ontologías mediante reutilización de recursos ontológicos.
- Escenario 4. Desarrollo de redes de ontologías mediante reutilización y reingeniería de recursos ontológicos.

²⁰<http://www.neon-project.org/nw/Welcome to the NeOn Project>

- Escenario 5. Desarrollo de redes de ontologías mediante reutilización y mezcla de recursos ontológicos.
- Escenario 6. Desarrollo de redes de ontologías mediante reutilización, mezcla y reingeniería de recursos ontológicos.
- Escenario 7. Desarrollo de redes de ontologías mediante reutilización de patrones de diseño ontológicos.
- Escenario 8. Desarrollo de redes de ontologías mediante reestructuración de recursos ontológicos.
- Escenario 9. Desarrollo de redes de ontologías mediante la localización de recursos ontológicos.

La Metodología Neón no sigue un flujo de trabajo rígido, por el contrario, sugiere una variedad de caminos para la construcción de redes de ontologías. Además, ofrece un kit de herramientas que incluye varios plug-in que apoyan las tareas del proceso de desarrollo. En la ilustración 9 podemos observar los escenarios para construcción de ontologías y redes ontológicas sobre los cuales trabaja la metodología Neón.

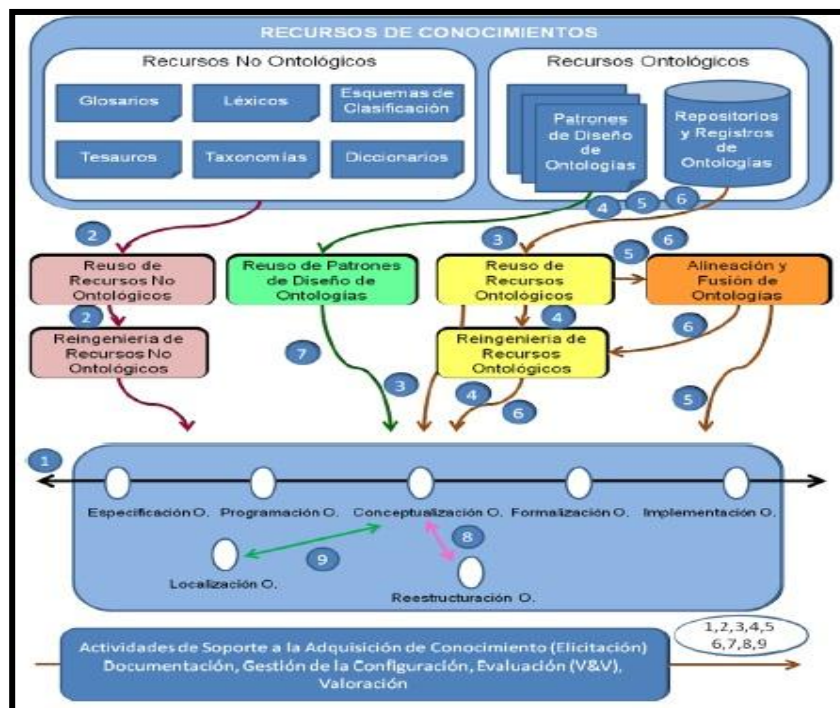


Ilustración 9 Escenarios de la Metodología Neón (Suárez, 2010)

1.10. Trabajos Relacionados en el ámbito de legislación.

Desde la década de 1940 aparecieron los primeros sistemas de información jurídica, en 1950 los sistemas expertos legales y en la actualidad también se está trabajando en temas jurídicos o gubernamentales que sean de acceso libre, cabe destacar que en diferentes países como: Chile, Holanda, Estados Unidos, Reino Unido, Australia, Austria, Canadá, Italia, Nueva Zelanda se están desarrollando proyectos de esta índole.(Cifuentes F. , 2011)

La cantidad creciente de datos legales gestionados por diferentes empresas jurídicas, bufetes de abogados y agencias de gobierno junto con los altos requisitos de calidad aplicables a la búsqueda de información legal-conocimiento, el descubrimiento y la gestión han llevado a la necesidad de desarrollar e implementar mejores herramientas de gestión de contenidos y métodos. (Casellas N. , 2010)

En la actualidad los datos que existen en el ámbito legal surgen de las diferentes actividades de los ciudadanos en internet, la información la podemos encontrar en: Foros de los ciudadanos, Plataformas de recolección de datos públicos y comentarios sobre ellos y Blogs de expertos jurídicos (Blawgs).

Algunos de los beneficios de aplicar los principios de Linked Open Data para la publicación de información en el ámbito de leyes son (Casellas et al., 2012):

- Mejor acceso (unificado, universal y sencillo) y comprensión del conocimiento legal desde el punto de vista de los ciudadanos y de las organizaciones.
- Mejorar la búsqueda y la gestión integrada de información y conocimiento.
- Permitir la compartición y reutilización de información legal entre aplicaciones, organizaciones y jurisdicciones.
- Reducción de la inversión de herramientas y sistemas y mejora de la cooperación e integración de organismos gubernamentales de cada país y entre países; esto es posible gracias al uso de vocabularios consensuados para la definición de términos y documentos legales.

Para obtener estas ventajas, primero se deben superar algunas cuestiones que dificultan el modelado y gestión de datos legales: la mayoría de los procesos judiciales se centran en documentos pero éstos están en diferente formato, medio, vocabulario, repositorio y lenguaje.

Algunos proyectos relacionados con Linked Legal Data:

Legislation.gov.uk fue lanzado por Reino Unido el cual permite a los ciudadanos acceder de forma pública a la información jurídica. Este proyecto abarca las cuatro jurisdicciones las cuales son: Inglaterra, Escocia, Wales e Irlanda del Norte.

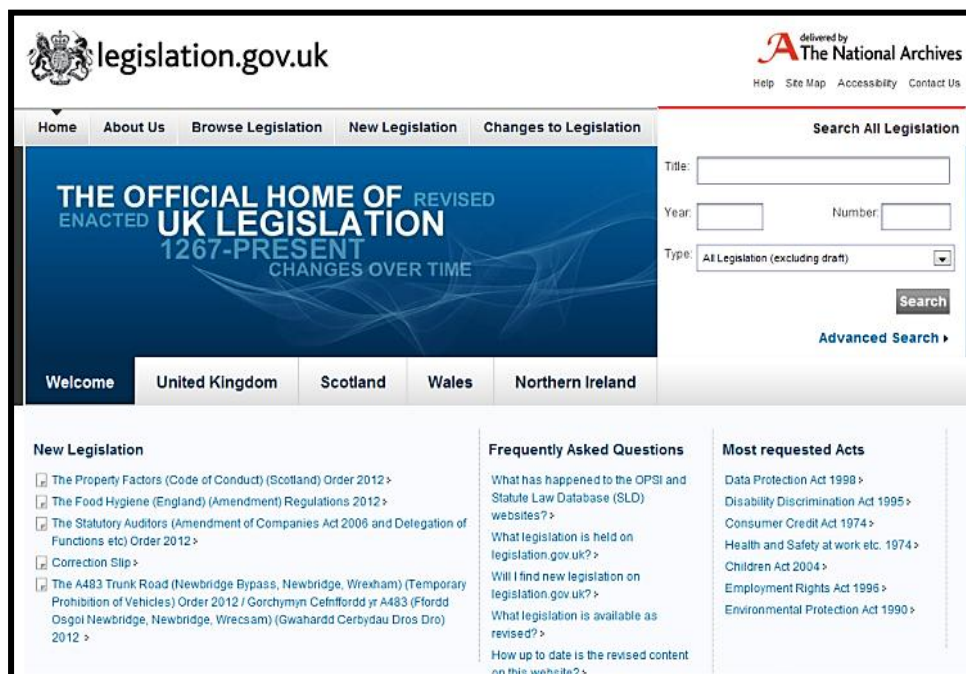


Ilustración 10 Portada de legislation.gov.uk

Según John Sheridan con la legislation.gov.uk se tenía dos objetivos los cuales son: ofrecer un servicio de alta calidad para las personas cuando necesiten consultar, citar. Y utilizar la legislación en línea y exponer el estatuto de Reino Unido; la finalidad es mostrar la legislación con el fin de contribuir a la Web de Datos. La legislación es difícil de entender ya que una ley puede tener diferentes secciones cada una con una fecha diferente de inicio. Lo que se trata con la legislation.gov.uk es tener una interfaz de usuario que permita que la ley sea clara. (Casellas, Vallbé, & Bruce, 2011)

Para usuarios más avanzados existe una línea de tiempo que puede ser activado para ver cómo la legislación ha cambiado y para navegar a través de una Ley en determinados puntos en el tiempo, incluidas las versiones futuras o potenciales. legislation.gov.uk es un sitio web atractivo, permitiendo un acceso sencillo y directo a la legislación, la gente puede ver Hechos enteros, o una sección en particular, ya sea en formato HTML o en versión en PDF.

En el proyecto se utilizaron algunos estándares abiertos como son: XML, XSLTs y XQueries . También utiliza software Open Source como: Orbeon Forms, Squid, and Apache.

Otra iniciativa proveniente del **gobierno del Reino Unido** ha puesto a disposición de la ciudadanía el portal data.gov.uk²¹ en el cual presentan a través de datos enlazados, información pública de interés ciudadano. En la actualidad cuentan con 9564 datasets sobre departamentos centrales del gobierno y otras entidades públicas y autoridades locales.

Este proyecto además de datos publicados, posee una infraestructura de soporte a datos enlazados, dentro de ello: Documentación desarrollada para el consumo de los datos, un Endpoint SPARQL para consultas, un buscador semántico, y otras aplicaciones de visualización de datos enlazados específicas para su dominio.



Ilustración 11 Portada de Data.Gov.uk

²¹<http://data.gov.uk>

Por otra parte, el **Instituto de Información Legal de la Universidad de Derecho de Cornell** tiene como objetivo tratar que la ley sea gratuita, accesible y comprensible. La Dra. Nuria Casellas participa en proyectos relacionados con Principios de Linked Open Data aplicados al código de regulaciones federales (CFR)²² el cual se ha publicado en semanticweb.com²³. LII²⁴ planea incorporar definiciones y vocabularios y la información del producto con la finalidad de que la búsqueda y recuperación de información mejore considerablemente.

La nueva versión de CFR tendrá nuevas características:

- Vínculos con la autoridad legal pertinente y reglamentación de expedientes que puedan afectar al usuario.
- Nuevas Funciones basadas en Web Semántica.
- A futuro en el CFR se podrá realizar las búsquedas por código de producto de las Naciones Unidas, la identificación y vinculación de la información relevante y una amplia variedad de Datos Vinculados.

Según (Casellas N. , Legally Linked: Linked Open Data Principles Applied To Code Of Federal Regulations, 2012) también se intenta conectar con los recursos externos que es el cuarto principio de Linked Data. Por ejemplo vincular el banco de drogas de datos abierto y base de datos de destino de drogas, que se ha transformado en RDF y está disponible como un punto final de SPARQL.

A través de la implementación de la extracción de la información mediante varias técnicas naturales de procesamiento de lenguaje: desarrollo de vocabulario SKOS, extracción de definiciones y la capacidad de relacionarse con secciones del CFR.

Con el desarrollo de este repositorio de datos RDF, la estructura y el contenido de los reglamentos permitirán crear aplicaciones basadas en Linked Data con el fin de mejorar la navegación, el descubrimiento y la agregación de la materia en el Código de Regulaciones Federales.

²²<http://www.law.cornell.edu/cfr/text>

²³http://semanticweb.com/legally-linked-linked-open-data-principles-applied-to-code-of-federal-regulations_b29197

²⁴Legal Information Institute

Estados Unidos es otro de los países que está a la vanguardia de la tecnología Linked Data o datos vinculados. El portal data.gov²⁵ contiene aproximadamente 6.400 millones de tripletas de datos gubernamentales abiertos. Para el cual han definido índices de documentos RDF que están disponibles para ser consultados y utilizados.



Ilustración 12 Portal de Data.Gov

Dentro del contexto legislativo, la Biblioteca del Congreso de los Estados Unidos posee una gran infraestructura de soporte a datos enlazados puesta a disposición de la ciudadanía. Algunos servicios ofrecidos son: cantidad de vocabularios, documentación y visualizaciones disponibles para el consumo de datos enlazados.

La Universidad de Amsterdam también incursiona en el ámbito de la legislación mediante Linked Open Data, el servidor de documentos Metalex. El cual proporciona identificadores persistentes, los metadatos se publica como datos RDF enlazados, utiliza vocabularios estándar como: Dublin Core, FOAF, Ontología Metalex.

²⁵<http://data.gov>

El contenido puede ser recuperado mediante la URI, los metadatos pueden ser accesibles a través de SPARQL. Los resultados que se desea obtener es asegurarse que la información más reciente estará siempre disponible, fácil de usar, entender y totalmente gratuita.

En el artículo denominado “Law and the Semantic Web, an Introduction”, explica que el uso de la tecnología de la web semántica en el ámbito jurídico, desempeña un papel fundamental.

Algunas de las tecnologías que se puede utilizar son RDF, XML, además existen varios tipos de aplicaciones que van del acceso de texto de la ley ya sea en línea o en CD. Existen aplicaciones que ayudan en la gestión administrativa de los bufetes de abogados, consulta de bases de datos jurídicas proporcionan un acceso rápido al contenido que permiten diferentes criterios de búsqueda tales como: tipo de fallo, la apelación, la fecha, el resumen, el número de fallo (Benjamins & Casanovas, 2012) .

En el artículo “Applying Human Language Technology to the Law” se encuentra información de avances generalizados en la web de distribución y procesamiento de la información textual legal, por ejemplo: casos, la legislación, las fuentes de información ciudadana, etc. Recientemente, un grupo creciente de investigación se ha ocupado de una serie de temas comunes tanto para el Equipo de Alto Nivel y las comunidades de Inteligencia Artificial y Derecho, incluyendo el razonamiento jurídico automatizado y argumentación, la recuperación de información semántica, clasificación de documentos, lógica de representaciones de lenguaje jurídico, los sistemas de diálogo, de redacción jurídica, el descubrimiento y la extracción de conocimientos jurídicos, lingüísticos basados en ontologías legales, entre otros.(Wyner, 2011).

Lo esencial de estos temas comunes es el uso de las técnicas del Equipo de Alto Nivel y herramientas para la automatización de conocimiento de la extracción de los textos jurídicos y para el procesamiento de lenguaje jurídico.

Otra iniciativa en el ámbito legislativo es el **Proyecto Europeo de Standardized Transparent Representations in order to Extend Legal Accessibility** denominado “Estrella”²⁶.

El objetivo es desarrollar y validar una plataforma abierta y basada en estándares que permite la administración pública para desarrollar e implementar soluciones integrales de gestión de conocimientos jurídicos, sin depender sobre productos patentados de vendedores particulares. Además apoyará, de manera integral, tanto en la gestión de documentos jurídicos y los sistemas basados en los conocimientos jurídicos, para proporcionar una solución completa para la mejora de la calidad y eficiencia de los procesos determinantes de la administración pública que requiere la aplicación de una legislación compleja y otras fuentes jurídicas (Estrella, 2008).

El proyecto Estrella intenta desarrollar un formato de intercambio de conocimientos jurídicos (LKIF), haciendo uso de las nuevas normas basadas en XML de la Web Semántica, así también RDF y OWL, e interfaces de programación de aplicaciones (API) para interactuar con el conocimiento legal.

Otra iniciativa **Norme in Rete (NIR)**²⁷ que intenta fomentar la transparencia y la accesibilidad de información jurídica para los ciudadanos comunes. Mediante el portal del proyecto se ofrece un punto de acceso unificado a toda la documentación legislativa publicada en los sitios web de los organismos institucionales italianos.

El Proyecto NIR tiene como objetivo proporcionar una herramienta que permita a los ciudadanos ejercer su derecho a la información y al mismo tiempo que permita a la institución cumplir su obligación de facilitar el acceso (ITTIG, 2010). Para estructurar los textos y normativa legislativa, NIR propone un esquema basado en definiciones de datos XML.

²⁶<http://www.estrellaproject.org/>

²⁷<http://www.ittig.cnr.it/Ricerca/UnitaEng.php?Id=40>

Akoma Ntoso²⁸ propone un conjunto de estándares, lenguajes y guías para la definición de documentos generados en los ámbitos: parlamentario, legislativo y documentos judiciales.

Akoma Ntoso busca facilitar el acceso a materiales legales primarios, obras y documentos parlamentarios poder judicial no es sólo una cuestión de dar acceso físico o en línea con ellos. "Acceso abierto", se describe y clasifica de manera uniforme y organizada para que el contenido se estructura en elementos significativos que pueden ser leídos y comprendidos por las aplicaciones de software, por lo que el contenido se hace "de lectura mecánica" y aplicaciones más sofisticadas de visualización en pantalla(Akoma Ntoso, 2011).

Akoma Ntoso cubre algunos aspectos como:

- Formato de documentos.- utiliza XML como lenguaje para definir la estructura y la sintaxis de documentos abiertos que son parte de: debates parlamentarios, sumarios, expedientes de comité, journals, reuniones del cuerpo de legislación, sentencias y otros.
- Modelo para el intercambio de documentos.- existen similitudes compartidas entre documentos del ámbito legal a pesar de: ser de diferentes tipos, tener diferentes propósitos y provenir de diferentes países; el modelo intenta identificar y describir éstas similitudes, con el fin de unificar y simplificar, los formatos y herramientas de software relacionados con la documentación legal.
- Otro objetivo que se pretende lograr es ofrecer un modelo de acceso abierto, centrado en: el uso de herramientas comunes para la generación de documentos (independientemente del tipo de proceso, país, idioma, y generación).Con el fin de crear repositorios, motores de búsqueda, herramientas de análisis, herramientas de comparación, etc.
- Esquema de datos.- soporta la especificación de diferentes tipos de documentos sobre: legislación, registros de debate, cuestiones

²⁸<http://www.akomantoso.org/><

parlamentarias, procedimientos judiciales y sentencias, etc; considera las características específicas relativas a estos tipos de documentación. Los documentos comparten las mismas estructuras básicas, las cuales han sido identificadas para establecer esta estructura común y que a su vez pueda adaptarse a las características de cada nación.

- Esquema de metadatos y ontologías.- la ontología sobre el dominio parlamentario cubre el conjunto básico de elementos de datos necesarios para la gestión y la recuperación de información oficial parlamentaria, legislativa y judicial. La ontología akoma Ntoso utiliza los valores y términos extraídos de tesauros del dominio legal, para mejorar las búsquedas que realizan profesionales del derecho. El namespace es el siguiente: “<http://www.akomantoso.org/2.0>”.

Ontologías Legales

El Instituto de Derecho de la Universidad de Cornell acerca del uso de ontologías en el ámbito del derecho y las leyes. Es importante destacar que las ontologías legales pueden ser distinguidas por los niveles de abstracción de las ideas que ellas representan.(Cornell University Law School, 2010)

Es así que se pueden encontrar ontologías generales (core ontologies, modelan conceptos generales y pueden servir para la comprensión de las leyes usadas en el dominio legal) o más específicas de un dominio determinado.

Las ontologías de dominio legal, por el contrario, se orientan a representar el conocimiento conceptual en relación a áreas específicas de la ley o dominios de práctica, y se construyen orientadas a una aplicación específica como: permitir la comunicación (vocabularios compartidos), o mejorar la indexación, búsqueda y recuperación de información jurídica.

Otras ontologías Core o generales de modelación en el ámbito legal, se citan:

- **LKIF-Core Ontology#²⁹**.- la ontología define un conjunto de aproximadamente 200 conceptos que van desde lo abstracto a concreto que se puede utilizar para establecer las bases para un sistema basado en el conocimiento legal. La ontología es desarrollada por el Consorcio

²⁹<http://www.estrellaproject.org/lkif-core/>

Estrella³⁰ y su base URI es: <http://www.estrellaproject.org/lkif-core/lkif-core.owl>.

- **LRI-Core Ontology#**³¹.- es una ontología que cubre los conceptos principales que son comunes al ámbito jurídico. Es utilizada en el proyecto del Centro Leibniz de Derecho de la Universidad de Amsterdam. LRI-CORE tiene las siguientes categorías: conceptos físicos, conceptos mentales, funciones, conceptos abstractos y términos de ocurrencias que se utilizan para hablar de casos o situaciones de manera genérica(Leibniz Center for Law, 2010).
- **DOLCE+CLO (Core Legal Ontology#)**³².- y la *Ontology of Fundamental Legal Concepts#* incluye referencias a los conceptos, Norma, Acto Legal y Persona Legal, y ciertos operadores sobre derechos como: prohibición, obligación y permiso.

Algunas de las ontologías propuestas en diferentes proyectos son:

CEN METALEX(CEN MetaLex, 2010) es un formato de intercambio, la intención no es de sustituir las normas específicas de cada jurisdicción y formatos específicos del proveedor en el proceso de publicaciones, sino imponer una visión estandarizada de documentos legales a los efectos del intercambio de información y la interoperabilidad en el contexto de desarrollo de software. CEN Metalex³³ ha propuesto la ontología denominada, [metalex-cen.owl](http://www.metalex.eu/metalex-cen.owl)³⁴.

METALEX define un mecanismo para la extensión del esquema, la adición de metadatos, cruzar referencias, documentos compuestos construcción y una convención de nomenclatura básica. El Namespace de XML Schema and OWL Ontology es: <http://www.metalex.eu/metalex/1.0>.

³⁰<http://www.estrellaproject.org/>

³¹<http://www.leibnizcenter.org/previous-projects/lricore>

³²<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.143.1048&rep=rep1&type=pdf>

³³<http://www.metalex.eu/>

³⁴<http://justinian.leibnizcenter.org/MetaLex/metalex-cen.owl>

Surge la iniciativa de Chile Datos que promueve prácticas y estándares de publicación de datos públicos. La idea de la organización es que las personas puedan acceder tanto a los datos como a aplicaciones para procesarlos.

La Biblioteca del Congreso Nacional³⁵ tiene avances en Linked Open Data, logrando el nivel 5 en la clasificación de 1 a cinco estrellas que se utiliza para evaluar el nivel de avance en la implementación del sistema, siendo www.bcn.cl la única institución estatal latinoamericana en conseguir este reconocimiento.

La BCN tiene a disposición un conjunto de ontologías modeladas de diferentes dominios de datos. Las ontologías son: Normas, Biografías, Geografía, Congreso.

Estas han sido modeladas utilizando RDFS y OWL, se encuentra documentadas en español e inglés utilizando las propiedades `rdfs:label` y `rdfs:comment`. La ontología está descrita en formato N3 como en RDF/XML, posiblemente también será descrita en RDFa. Para la construcción de las ontologías la Biblioteca del Consejo Nacional de Chile ha reutilizado diferentes vocabularios como: FOAF(Friend of a Friend:), SKOS(Simple Knowledge Organization System), DC(Dublin Core:), DBPEDIA-OWL(DBPedia), GN(Geonames), ORG(Organization).

En la ilustración 13 se muestra las principales clases y sus relaciones de la Ontología de Normas de la Biblioteca del Congreso Nacional de Chile.

³⁵<http://datos.bcn.cl/portal/ontologias/>

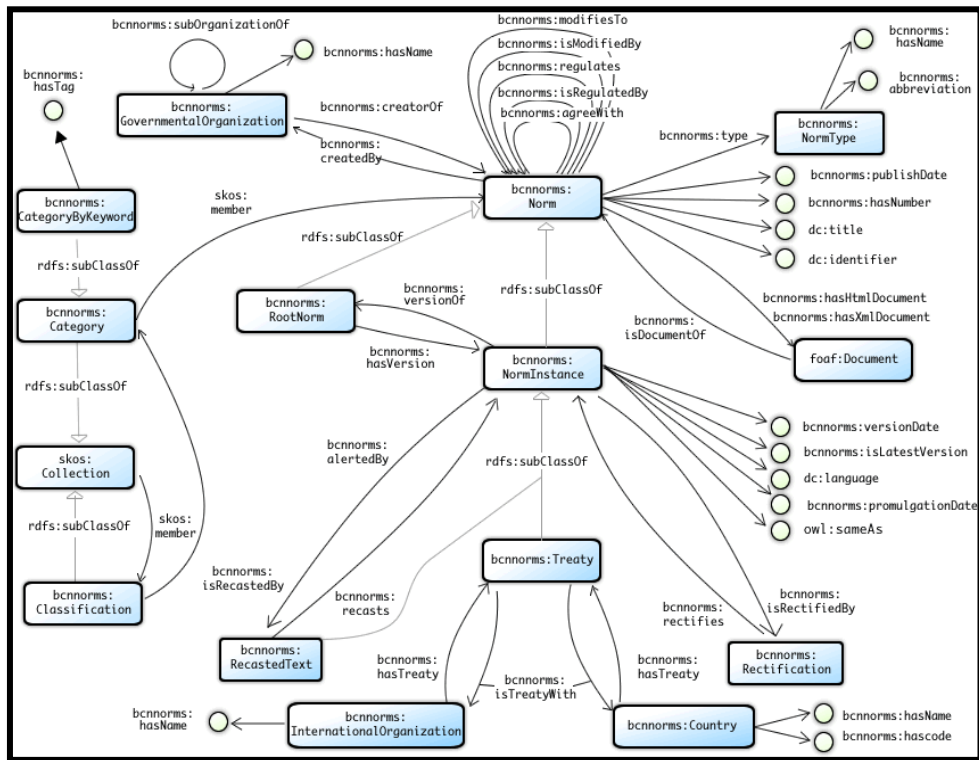


Ilustración 13 Ontología de normas del Congreso de Chile (Biblioteca del Congreso Nacional de Chile, 2011)

Mediante el capítulo del estado del arte se comprendió varios conceptos tales como: Linked Data, Web Semántica, RDF, RDFS JENA, SPARQL los cuales ayudaran al desarrollo de la investigación, también se analizaron diferentes herramientas de almacenamiento para RDF como son: AllegroGraph, Sesame, Virtuoso y 4Store; de las diferentes ontologías que se analizaron la que más se relaciona con el ámbito de Legislación de Educación Superior es la Ontología de Normas de la Biblioteca Nacional de Chile esta ontología describe las normas legislativas de Chile y sus relaciones en el contexto legislativo, actualmente en el ámbito de Legislación Universitaria no existe un vocabulario genérico por lo cual se surge la necesidad de crear un nuevo vocabulario que cubra el dominio del proyecto con la finalidad de resolver las interrogantes propuestas.

**2. PUBLICACIÓN DE DATOS ENLAZADOS EN EL ÁMBITO DE LEGISLACIÓN
UNIVERSITARIA**

INTRODUCCIÓN

El presente capítulo contiene: modelo del proyecto y proceso de publicación de datos enlazados el cual se encuentra dividido en seis fases: Identificación y Extracción de datos, Desarrollo del Vocabulario, Limpieza de Datos, Publicación de Datos RDF y Enlazado de datos las mismas que se detallan en el desarrollo del capítulo.

2.1. Modelo del Proyecto.

El modelo del proyecto con el cual se trabajó para el proceso de publicación de datos enlazados se lo puede observar en la ilustración 14. En la fase 1 se realiza la identificación de datos estructurados y no estructurados sobre Legislación de Educación Superior los cuales se encuentra en diferentes formatos: pdf, word y páginas web. Luego de identificar la información con la ayuda del personal de gestión legal e informático se determinó los metadatos a extraer los cuales fueron almacenados en hojas de Excel.

En la fase 2 se desarrolla el vocabulario para lo cual se utiliza parte de la metodología NEON para determinar los requisitos con los que se abarca el dominio del proyecto, también se reutiliza vocabularios genéricos que nos permitan construir el vocabulario.

En la fase 3 se realiza la limpieza de datos utilizando la herramienta Google Refine permite trabajar con grandes volúmenes, esta herramienta genera datos de mejor calidad; luego de realizar este proceso se almacena los datos en la base MYSQL denominada Leyes. En la fase 4 se realizó la generación de datos RDF para lo cual se utilizó el api de Jena y para acceder a los datos se realiza una conexión hacia la base de datos MYSQL y se procede a instanciar el vocabulario; luego de esto se obtiene un archivo denominado RDF.

En la fase 5 se realizó la publicación de datos RDF en el servidor Virtuoso para lo cual se carga el archivo generado en la fase 4 luego a través del ENDPOINT mediante consultas SPARQL planteadas al inicio de la investigación.

Y en la fase 6 se realiza el enlazado de datos con diferentes fuentes.

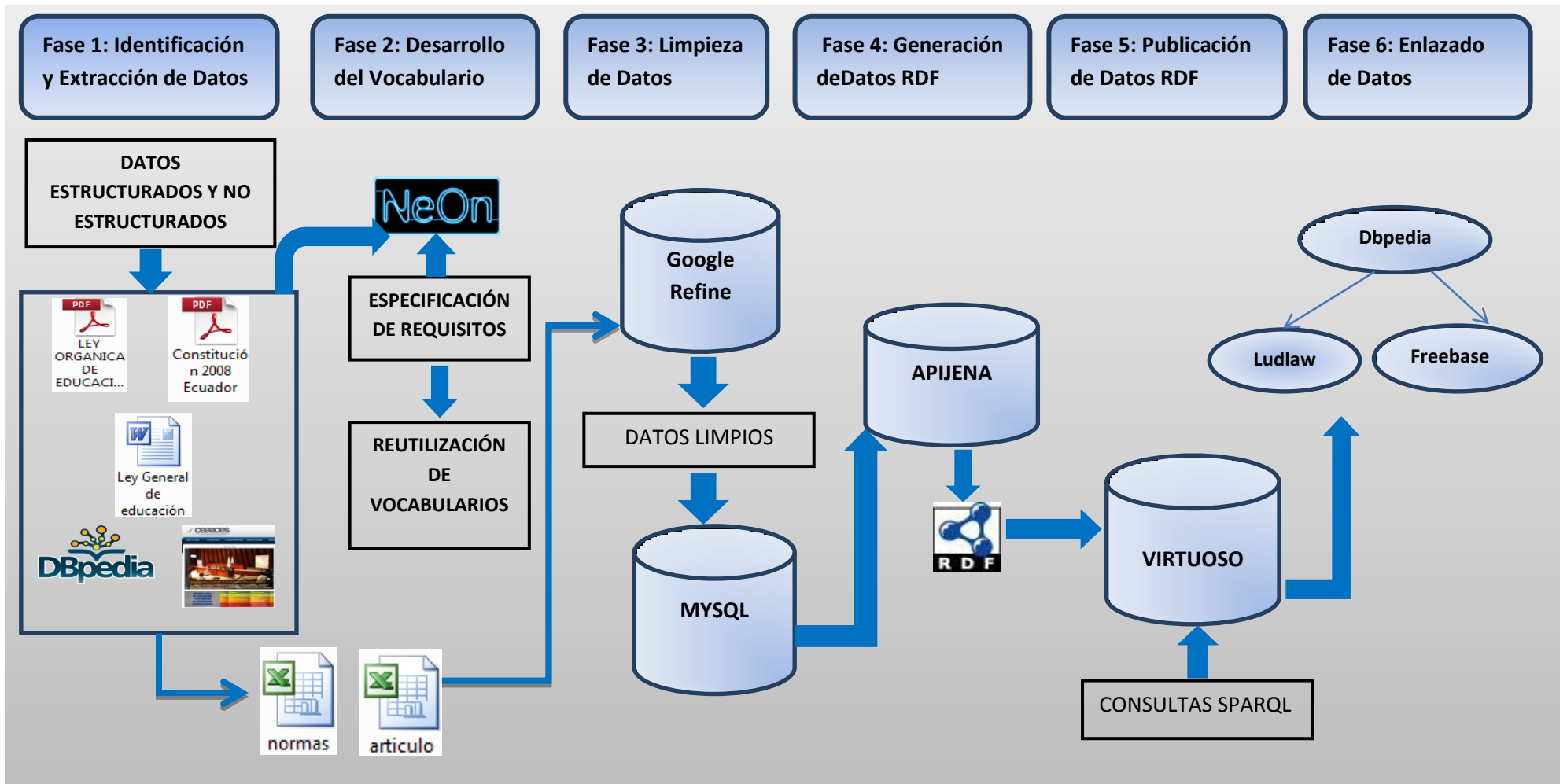


Ilustración 14 Modelo del Proyecto

FASE 1

2.2. Identificación y Extracción de Datos.

2.2.1. Identificación de las fuentes de datos.

Los datos sobre Legislación de Educación Superior comprende: normas, universidades, países y organismos de educación los cuales fueron recolectados de diferentes fuentes de datos como son: pdf, word, Dbpedia y páginas web. En la actualidad esta información se encuentra dispersa y en diferentes formatos por ende no se cuenta con información sistematizada, completa y confiable lo cual dificulta la extracción automática de la información.

Para ello se utilizará los beneficios que brindan las nuevas tecnologías y su forma de organización como es la Web Semántica.

A continuación se describen las características de las fuentes de datos seleccionadas:

- **Normas**

De las normas de educación superior seleccionadas en la Tabla 6 se puede observar las siguientes características:

Tabla 5 Resumen de características de las normas

Países	Total: 23 países Incluyendo: Argentina, Brasil, Chile, Colombia Costa Rica, Cuba, Ecuador, El Salvador, Guatemala, Honduras, México, Nicaragua, Panamá, Paraguay, Perú, Puerto Rico, República Dominicana, Uruguay, Venezuela
Tipos de Normas de Educación Superior	Constitución , Leyes, Reglamentos, Resolución, Estatuto, Decretos, Acuerdo
Lenguajes	Español Inglés Portugués Francés
Formatos de archivos	.PDF .DOC

- **Universidades**

La información de Universidades se la obtuvo desde DBpedia, ya que este Dataset permite acceder a la información de forma libre. En la ilustración 13 podemos observar datos de la universidad como: abstract, canciller, tipo de universidad, motto, país al que pertenece, dirección estos son algunos datos sobre la Universidad Técnica Particular de Loja.

About: Universidad Técnica Particular de Loja
 An Entity of Type: [Private university](#), from Named Graph: <http://dbpedia.org>, within Data Space: [dbpedia.org](#)

La Universidad Técnica Particular de Loja es una institución autónoma, con finalidad social y pública, pudiendo impartir enseñanza, desarrollar investigaciones con libertad científica – administrativa, y participar en los planes de desarrollo del país, otorgar, reconocer y revalidar grados académicos y títulos profesionales, y en general, realizar las actividades propias para la consecución de sus fines.

Property	Value
dbpedia-owl:abstract	<ul style="list-style-type: none"> La Universidad Técnica Particular de Loja (UTPL) was officially recognized by the Ecuadorian State according to the executive decree No. 646 published in the official record No. 217 on May 5, 1971. With this the UTPL becomes an autonomous corporate organization that is protected by the "Modus Vivendi" document formalized between the Saint Principal Office and the Ecuadorian State, keeping in mind the norms of the Church in its organization and governance. It provides on-campus and distance education. The latter dates back to 1976, which makes it a pioneer in Latin America. La Universidad Técnica Particular de Loja es una institución autónoma, con finalidad social y pública, pudiendo impartir enseñanza, desarrollar investigaciones con libertad científica – administrativa, y participar en los planes de desarrollo del país, otorgar, reconocer y revalidar grados académicos y títulos profesionales, y en general, realizar las actividades propias para la consecución de sus fines. Desde su fundación hace 35 años se evidencia el ascenso de una pequeña universidad que se inició con tres facultades: Economía, Ingeniería Civil e Industrias Agropecuarias, en las instalaciones del Instituto Técnico Superior "Daniel Alvaréz Burneo", hasta la institución actual que tiene dos modalidades de estudio, clásica o presencial y la modalidad Abierta o de Estudios a Distancia, donde estudian más de 22 mil personas en 25 carreras.
dbpedia-owl:chancellor	<ul style="list-style-type: none"> dbpedia:Ph.D._José_Barbosa_Corbacho
dbpedia-owl:city	<ul style="list-style-type: none"> dbpedia:Loja_Ecuador
dbpedia-owl:country	<ul style="list-style-type: none"> dbpedia:Ecuador
dbpedia-owl:motto	Memento Ascendere Semper (It's always time to advance), Decide ser más (Decide to be more)
dbpedia-owl:numberOfUndergraduateStudents	20000 (xsd:integer)
dbpedia-owl:province	<ul style="list-style-type: none"> dbpedia:Loja_Province
dbpedia-owl:type	<ul style="list-style-type: none"> dbpedia:Private_university
dbpedia-owl:wikiPageExternalLink	<ul style="list-style-type: none"> http://www.utpl.edu.ec/lepac http://www.utpl.edu.ec/ocor http://www.utpl.edu.ec/openutpl http://www.utpl.edu.ec/lefed http://www.confedec.com http://www.infosec.org/Intro.html http://www.umce.edu http://www.cic.edu.ec/ http://repositorio.utpl.edu.ec http://www.aesad.org/index1.html http://www.cread.org http://www.infuseduca.org http://www.utpl.edu.ec/lealed http://www.utpl.edu.ec/lelam http://www.utpl.edu.ec/
dbpprop:address	San Cayetano Alto
dbpprop:chancellor	dbpedia:Ph.D._José_Barbosa_Corbacho
dbpprop:city	dbpedia:Loja_Ecuador
dbpprop:country	dbpedia:Ecuador

Ilustración 15 Metadatos en DBpedia sobre la UTPL

- **Organismos de Educación Superior**

El dominio también abarca información sobre organismos de Educación Superior para lo cual se accedió a cada sitio oficial de los organismos, y se obtuvo los siguientes metadatos: nombre del organismo, tipo de organismo, siglas, fecha de fundación y homepage.



Ilustración 16 Sitio Oficial del organismo CEAACES

2.2.2. Extracción de los datos.

Para obtener la información sobre las normas de educación superior se contó con el apoyo de un grupo de abogados de la UTPL los cuales son los expertos en el dominio sobre legislación, conjuntamente con el equipo de sistemas se determinaron los metadatos a extraer sobre las normas.

Los metadatos que se extrajeron de la norma son:

- Título
- Fecha de promulgación
- Fecha de publicación
- Tipo de norma
- Jerarquía de la norma
- Idioma
- Presidencia
- Creador de la norma
- Lugar de publicación de la norma
- Número de la norma
- Url de la norma

En la ilustración 17 se puede observar algunos de los metadatos extraídos de las normas:

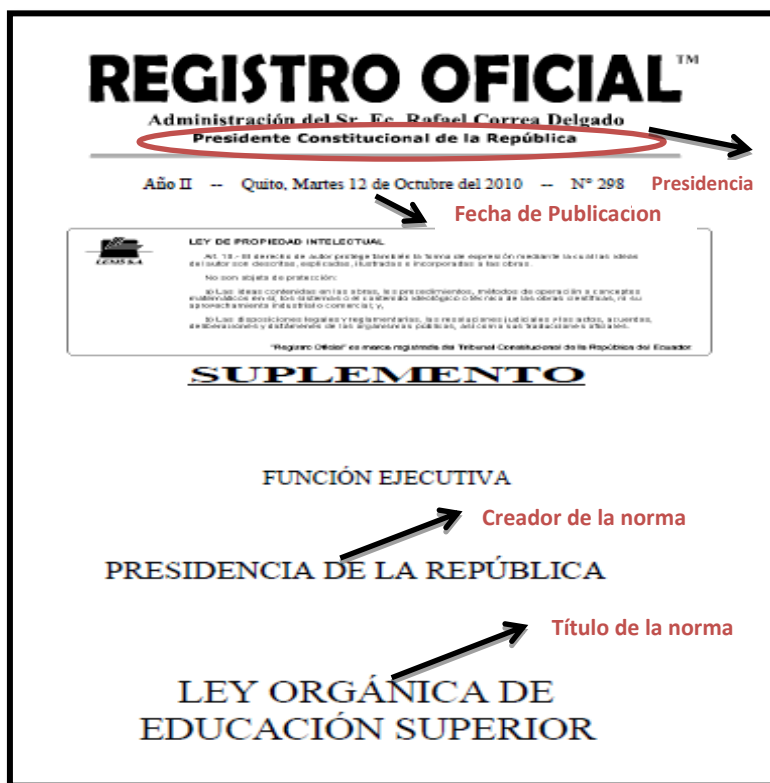


Ilustración 17 Metadatos de la Norma de Educación Superior

La información de las normas fue extraída mediante proceso manual, debido a que a la información se encuentra en diferentes formatos lo cual dificultó que la extracción se la realice mediante proceso automático.

Para extraer la información sobre las normas se realizó previamente una selección de los tipos de normas de educación superior como: constitución y ley de los cuales se extraerá los metadatos mencionados anteriormente.

Estos metadatos se los almacenó en hojas de Excel ya que es una herramienta que cuenta con diferentes funcionalidades, para el tratamiento de información.

En la tabla 6 se puede observar algunos de los metadatos que se extrajo de las normas de educación superior:

Tabla 6: Metadatos de las normas

TITULO	FECHA_P ROMULGA CION	FECHA_P UBLICACI ON	TIPO_NORM A	JERARQUI A_NORMA	IDIOMA	PRESIDENC IA	CREADOR
Ley Orgánica de Educación Superior	04/08/2010	12/10/2010	Ley	Nivel 3: Legislación Ordinaria	Español	Rafael Correa Delgado	Función Legislativa
Constituci ón de la República De Ecuador	30/11/2007	20/10/2008	Constitución	Nivel 1: Constitució n	Español	Rafael Correa Delgado	Función Legislativa
Ley de Educación Superior	07/08/1995	10/08/1995	Ley	Nivel 3: Legislación Ordinaria	Español	Carlos Saúl Menem	Función Legislativa

Luego de obtener los metadatos de las normas, se realizó un análisis del contenido para poder extraer los artículos de interés y clasificarlos por categorías para lo cual se revisó cada uno de los artículos que contienen la norma y se escogió los similares entre las demás normas y se los agrupo por algunas de las categorías como son: derecho a la educación, fines de la educación, autonomía universitaria, financiamiento universitario, autoridades académicas.

También se obtuvo los tags o palabras claves mediante las cuales se conocerá cual es enfoque y contenido del artículo.

En la tabla 7 se puede observar los datos que se obtuvo de cada uno de los artículos que contienen las normas:

Tabla 7: Datos de los artículos

NUM_ART	DESCRIPCIÓN	CATEGORÍA	TAGS
Art. 4	Derecho a la Educación Superior.- El derecho a la educación superior consiste en el ejercicio efectivo de la igualdad de oportunidades, en función de los méritos respectivos, a fin de acceder a una formación académica y profesional con producción de conocimiento pertinente y de excelencia.	Derechos Educación	educación superior, igualdad de oportunidad, formación académica, excelencia, conocimiento pertinente, producción conocimiento
Art. 52	Los estatutos de las instituciones universitarias nacionales deben prever sus órganos de gobierno, tanto colegiados como unipersonales, así como composición y atribuciones. Los órganos colegiados tendrán básicamente funciones normativas generales, de definición de políticas y de control en sus respectivos ámbitos, en tanto los unipersonales tendrán funciones ejecutivas.	Autoridades Académicas	estatutos, instituciones universitarias, definición de políticas

Universidades

La información de las universidades de Latinoamérica se la obtuvo desde DBPEDIA, para ello se utilizó la siguiente consulta SPARQL para obtener datos de las universidades de los países latinoamericanos:

```

SELECT DISTINCT
?s
?ontology_affiliation
?ontology_abstract
?ontology_campus
?ontology_chairman
WHERE {
?s a <http://dbpedia.org/ontology/EducationalInstitution>,
<http://dbpedia.org/ontology/University> .
?s <http://dbpedia.org/ontology/country> ?country .
?country <http://purl.org/dc/terms/subject>< South_American_countries> .
OPTIONAL {?s<http://dbpedia.org/ontology/affiliation> ?ontology_affiliation .}
OPTIONAL {?s<http://dbpedia.org/ontology/abstract> ?ontology_abstract .}
OPTIONAL {?s<http://dbpedia.org/ontology/campus> ?ontology_campus .}
OPTIONAL {?s<http://dbpedia.org/ontology/chairman> ?ontology_chairman .}
?property_long .}
OPTIONAL {?s<http://dbpedia.org/ontology/logo> ?ontology_logo .}
FILTER ( langMatches(lang(?ontology_abstract), "es") ||
langMatches(lang(?ontology_abstract), "en") )
FILTER ( langMatches(lang(?ontology_motto), "es") ||
langMatches(lang(?ontology_motto), "en") )
}

LIMIT 1000

```

Organismos de Educacion Superior

Los metadatos sobre los organismos reguladores de la de educación superior se la obtuvo de diferentes sitios oficiales de cada organismo, algunos metadatos que se extrajeron se puede observar en la tabla 8:

Tabla 8: Datos de organismos

Nombre del organismo	Tipo organismo	Siglas	Fundación	Homepage
Consejo de Evaluación, Acreditación y Aseguramiento de la Calidad de la Educación Superior	Organismos que rigen el Sistema de Educación Superior	CEAAC ES	2010	http://www.ceaaces.gob.ec/
Consejo de Educación Superior	Organismos que rigen el Sistema de Educación Superior	CES	2010	http://www.ces.gob.ec/

FASE 2

2.3. Desarrollo del Vocabulario.

Con el fin de obtener cuales son las especificaciones para el desarrollo del vocabulario se utilizó parte de la metodología Neón(Suárez-Figueroa, 2008), de la cual se enfoca principalmente en el escenario 1 que es la especificación de los requisitos con la finalidad de dar respuesta a las siguientes interrogantes:

- ¿Para qué se construye la ontología?
- ¿Cuáles son los usos previstos de la misma?
- ¿Quiénes son los usuarios finales?
- ¿Cuáles son las necesidades que la ontología debe satisfacer?

Como resultado se obtendrá el Documento de Especificación de Requisitos de la Ontología (DERO).

1 Propósito

Proporcionar un modelo de conocimiento que represente la información sobre Legislación de Educación Superior, con la finalidad de dar a conocer cuales son las normas que rigen el sistema de educación superior de cada país de Latinoamérica.

2 Alcance, Ámbito

El alcance de la ontología sobre Legislación de Educación Superior en Latinoamérica abarca la siguiente información:

- Las Normas de Educación Superior que se analizarán son: leyes y constituciones de los cuales se extraerá varios metadatos de cada una de las normas. Cada una de las normas están compuestas por varios artículos por ende se analizará previamente y se obtendrá los artículos de mayor relevancia respecto a la Educación Superior puede ser: autonomía universitaria, derecho a la educación entre otros.
 - Sobre los organismos de educación Superior existen diferentes tipos: de evaluación y acreditación de la calidad, de Coordinación y consulta y organismos reguladores los cuales son los encargados
-

de mejorar la educación superior en cada país.

- La información sobre universidades se obtendrá desde Dbpedia.
- De los países de Latinoamérica se escogieron: Argentina, Ecuador y México de los cuales se obtuvo información de las normas que rigen el sistema de educación superior.

3 Nivel de formalidad

La ontología se ha modelado en CMAP TOOLS COE y el lenguaje de implementación es OWL.

4 Usuarios Potenciales

Los usuarios previstos para la ontología son:

- Estudiantes o profesionales en formación de educación superior.
- Profesores de educación superior encargados de participar el conocimiento.
- Instituciones de Educación Superior: forman parte las Universidades, Institutos y Colegios Universitarios.
- Organismos de Educación Superior: existen varios tipos de organismos tales como: organismo reguladores, organismos de evaluación y Acreditación de la calidad, organismos de coordinación y consulta los cuales son encargados de regular y mejorar la calidad de Educación Superior.
- Gobierno: autoridad que dirige, controla y administra las instituciones del Estado o creadoras de las normas de educación superior.

5 Usos Potenciales

- Conocer cuáles son las normas de tipo ley que rigen el sistema de Educación Superior.
- Conocer cuáles son las normas que mencionan sobre el derecho a la Educación Superior.
- Conocer el nombre del presidente que en su administración se crearon las normas de Educación Superior.
- Conocer cuáles son los artículos que mencionan sobre el personal académico.
- Buscar información sobre normas de Educación Superior con el fin de mejorar la calidad día a día.
- Conocer las instituciones de Educación Superior que forman parte

del sistema de Educación.

- Conocer que tipo de normas rigen la Educación Superior en los países de Latinoamérica.
- Conocer que países en su normativa mencionan sobre el derecho a la Autonomía Universitaria.
- Consultas relacionadas sobre los organismos reguladores de la Educación Superior en cada país de Latinoamérica(Tipo de organismo, fecha de fundación, descripción, siglas, homepage)
- Consultas sobre las normas que mencionan en su normativa sobre organismos de Educación Superior.
- Saber que organismos son los encargados de la coordinación y consulta de la Educación Superior.

6 Preguntas de competencia

- ¿Cuáles son las normas de educación superior y su fecha de publicación?
- ¿Cuáles son las normas de educación superior que son de tipo de ley?
- ¿Cuáles es el nombre del presidente que en su administración se crearon las normas de educación superior de Argentina?
- ¿Cuáles son las normas que en el articulado mencionado sobre derecho de educación?
- ¿Qué normas X regulan el sistema de educación Superior en Ecuador?
- ¿Cuáles son los países que en su normativa mencionan sobre el derecho de la Autonomía Universitaria?
- ¿Cuáles son las normas que en su articulado mencionan sobre el financiamiento universitario?
- ¿Cuáles son las universidades que son parte del sistema de educación superior de Ecuador?
- ¿En qué normas se definen los organismos de educación superior?
- ¿Cuáles son los organismos que regulan la educación superior en el país de Ecuador?
- ¿Cuáles son los artículos que mencionan acerca del personal académico en las normas de educación superior?
- ¿Cuáles son los artículos que mencionan acerca del personal académico en las normas de educación superior?

7 Glosario de términos

En la tabla 5 se definen los términos que forman parte del dominio sobre Legislación de Educación Superior de Latinoamérica, se definen los conceptos y las propiedades que se utilizarán

Tabla 9 Glosario de Términos

CONCEPTOS	DESCRIPCIÓN	TIPO	SINÓNIMOS
Norma	Contiene información sobre las normas que rigen la educación superior en los diferentes países.	Concepto	Normativa
Organismo	La clase Organismo contiene información de los organismos que rigen el sistema de educación superior.	Concepto	Organización
Artículo	Contiene cada uno de los artículos como: número, descripción, categoría y tags.	Concepto	
Universidad	La universidad es una institución educativa de tercer nivel.	Concepto	Institución
País	Contiene los países.	Concepto	Estado
Sistema Educación Superior	Se refiere al tipo de Sistema de Educación que tiene cada país.	Concepto	
Título	Es el título o nombre que se da a la norma.	Propiedad	Nombre
Fecha Promulgación	Es la formalización de la norma que generalmente la hace el Ejecutivo (el presidente), es decir es la fecha de creación de la norma.	Propiedad	
Fecha Publicación	La publicación de la norma es cuando aparece en un medio formal y desde donde se entiende que es conocido por todos los habitantes.	Propiedad	
Tipo Norma	Describe un tipo de norma en el contexto normativo de la educación superior.	Propiedad	
Jerarquía Norma	El tipo de jerarquía de las normas según la Pirámide de Kelsen: Nivel 1: Constitución Nivel 2: Convenios internacionales Nivel 3: Legislación Ordinaria: Códigos en General Nivel 4: Legislación Secundaria: Decretos, reglamentos, estatutos, resoluciones	Propiedad	
Lenguaje	Define el idioma de la norma.	Propiedad	Idioma
Presidencia	Nombre del presidente que en su administración fue creada las normas de educación superior.	Propiedad	Gobierno
Creador	La entidad encargada principalmente de crear la norma.	Propiedad	
Nombre Registro	Nombre asignado al registro que contiene y publica las normas.	Propiedad	
Numero Registro	Numero asignado al registro oficial.	Propiedad	

Url Norma	Es la dirección online donde se encuentra almacenada la norma.	Propiedad	
Nombre País	Es el nombre de cada país.	Propiedad	
Uri País	Es el uri Dbpedia del país.	Propiedad	
Establece	Las normas establecen a los organismos de educación Superior.	Propiedad	
Nombre Organismo	Nombre del organismo.	Propiedad	
Descripción Organismo	Una reseña del organismo.	Propiedad	
Siglas Organismo	Abreviatura del nombre del organismo.	Propiedad	
Fecha Fundación	Es la fecha en que el organismo fue fundado.	Propiedad	
Homepage	Una página web para algo.	Propiedad	
Tipo Organismo	Clasificación de los organismos.	Propiedad	
Logo Organismo	Elemento grafico que identifica a un organismo.	Propiedad	
Número Artículo	Parte numerada de un documento oficial.	Propiedad	
Texto Artículo	Es el texto que define a un artículo.	Propiedad	
Categoría	Es una síntesis de lo que contiene el artículo.	Propiedad	
Tag	Palabras que describen un tópico.	Propiedad	Palabras Claves
Nombre Universidad	Nombre de la universidad.	Propiedad	
Uri Universidad	Es el uri de la Universidad en Dbpedia.	Propiedad	
Compuesta de	La norma está compuesta por artículos.	Propiedad	Formada por
Is parte de	Los artículos son parte de una norma.	Propiedad	
Tiene Sistema	Cada país tiene un Sistema de Educación.	Propiedad	

Con la finalidad de observar de mejor manera los datos sobre Legislación de Educación Superior, se utilizó la herramienta CMAPTOOLS COE³⁶ ya que es una herramienta para el modelamiento.

³⁶<http://www.ihmc.us/groups/coe/>

En la ilustración 18 se puede observar el primer esquema del vocabulario LUDLAW que representa los conceptos sobre la legislación de Educación Superior, algunos de los conceptos a representar son: Norma, Organismo, Sistema de Evaluación de Acreditación, Artículo, Normativa Internacional, Organización Internacional, Fines Principios y Derechos Humanos, cada concepto contiene sus propiedades.

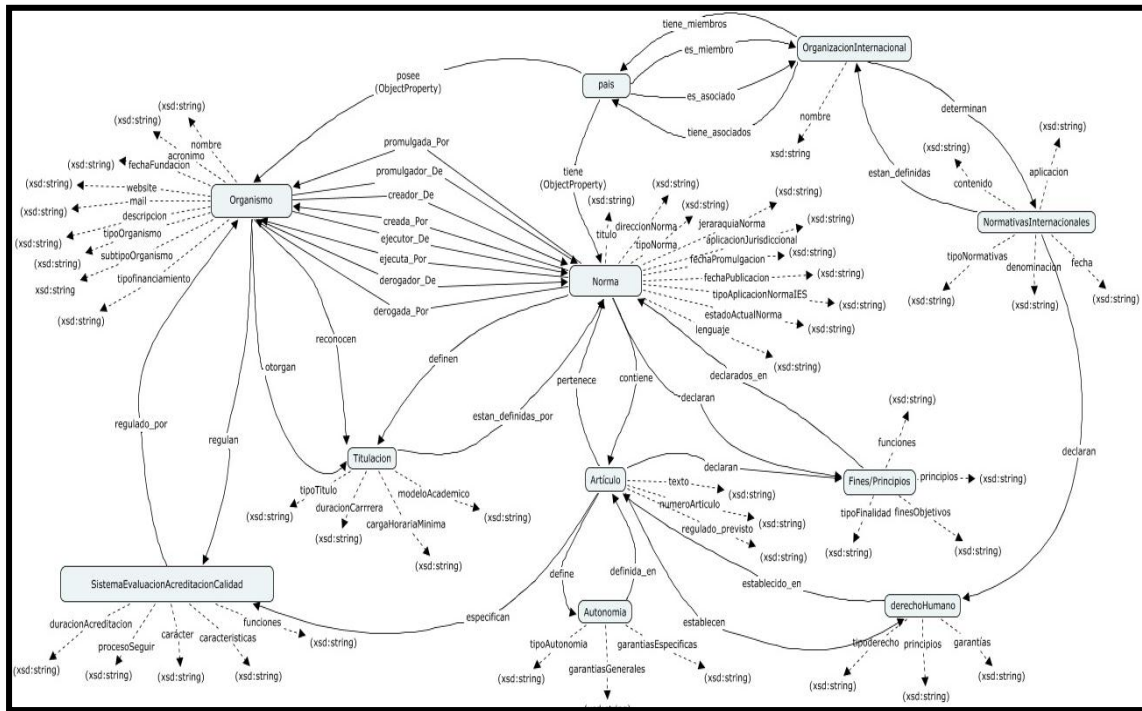


Ilustración 18 Esquema Preliminar LUDLAW

Debido a que la información hacer modelada estaba muy dispersa y extensa surge la necesidad de realizar algunos cambios: Se eliminó los conceptos Normativa Internacional y Organismos internacionales y algunos conceptos se los cambio a propiedades para representar de mejor manera la información.

En la ilustración 19 podemos observar el modelo con el que se trabajó para la construcción del vocabulario LUDLAW:

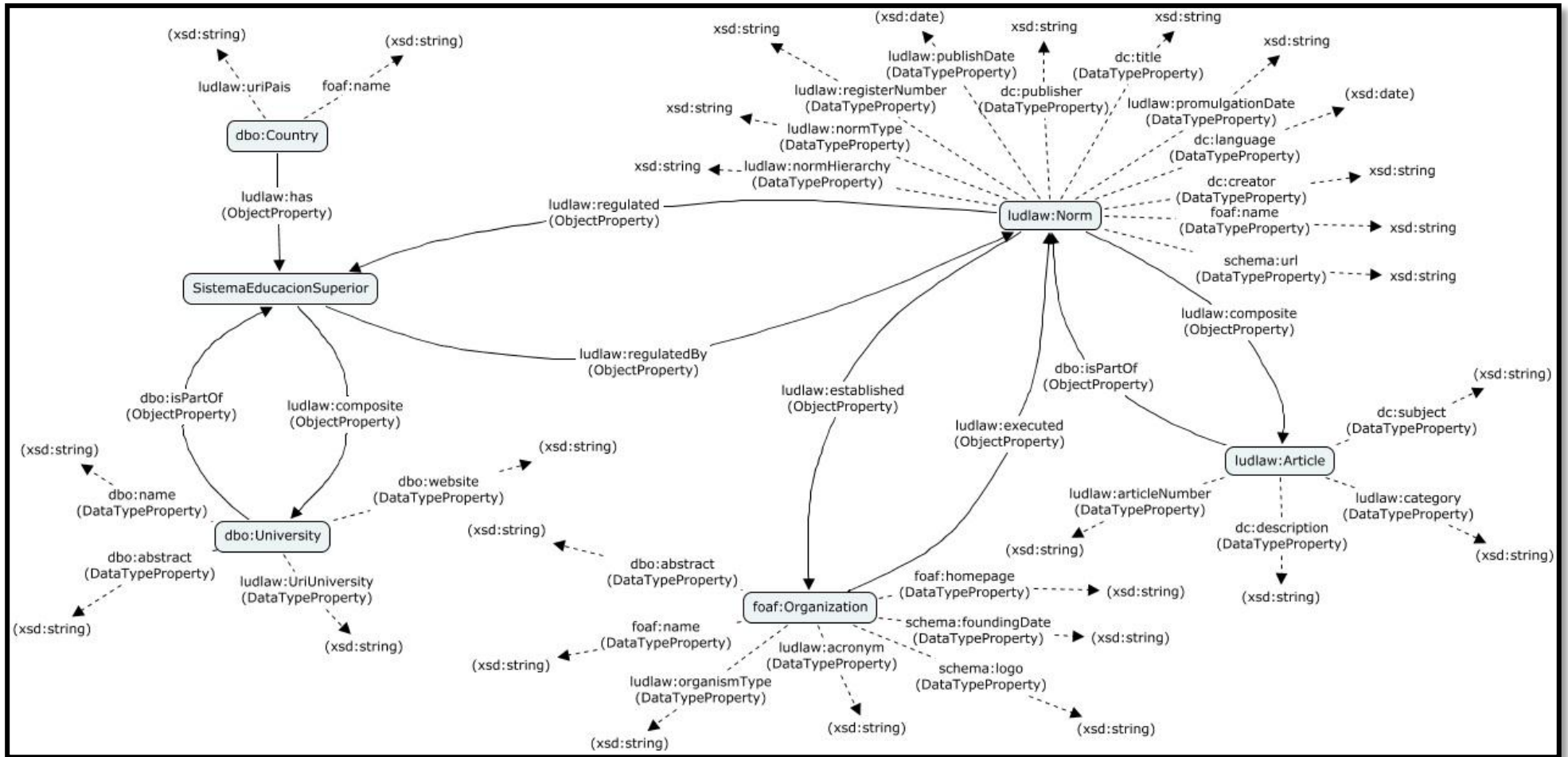


Ilustración 19 Modelo de Legislación de Educación Superior

Luego de haber identificado cuales son las especificaciones para la creación del vocabulario sobre Legislación de Educación Superior se procede a la construcción, para ello se definió las clases y propiedades que abarca el ámbito del proyecto.

Se reutilizará vocabularios existentes, la reutilización se constituye en un principio fundamental Linked Open Data, esto con la finalidad de que si ya existen los elementos en otros vocabularios no sean definidos nuevamente, mediante la reutilización se logra interoperabilidad semántica.

De acuerdo a la información a representar sobre Legislación de Educación Superior es posible utilizar los siguientes vocabularios:

- Dublin Core se utilizó para representar los metadatos de las normas.
- FOAF para definir a personas u organismos.
- DBpedia para obtener información sobre las universidades de Latinoamérica que forman parte del sistema de educación superior.
- AIISe utilizó para representar la información sobre los organismos de educación y para las universidades.
- BCNNORMS es una ontología sobre normas y sus relaciones en el contexto legislativo. Se utiliza para representar parte de las normas de educación superior.

Mediante la reutilización de vocabularios se intenta obtener un vocabulario consensuado sobre Legislación de educación Superior el cual contiene información sobre: Normas, Sistema de Educación, Organismo de Educación Superior, universidades, artículos. No todas las clases y propiedades de la información están definidas en los vocabularios antes mencionados, por ello se procede a la creación de un vocabulario que permita representar las clases y propiedades faltantes que abarque el dominio a representar.

El vocabulario se denominará **LUDLAW** y su URI es:<http://utpl.edu.ec/educacion/leyes/ludlaw/>.

Tabla 10 Vocabularios Reutilizados

CONCEPTOS	DESCRIPCIÓN	TIPO	VOCABULARIO
Norma	Contiene información sobre las normas que rigen la educación superior en los diferentes países.	Concepto	LUDLAW
Organismo	La clase Organismo contiene información de los organismos que rigen el sistema de educación superior.	Concepto	AIISO
Artículo	Contiene cada uno de los artículos como: número, descripción, categoría y tags.	Concepto	LUDLAW
Universidad	La universidad es una institución educativa de tercer nivel.	Concepto	AIISO
País	Contiene los países.	Concepto	LUDLAW
Sistema Educación Superior	Se refiere al tipo de Sistema de Educación que tiene cada país.	Concepto	LUDLAW
Título	Es el título o nombre que se da a la norma.	Propiedad	DC
Fecha Promulgación	Es la formalización de la norma que generalmente la hace el Ejecutivo (el presidente), es decir es la fecha de creación de la norma.	Propiedad	LUDLAW
Fecha Publicación	La publicación de la norma es cuando aparece en un medio formal y desde donde se entiende que es conocido por todos los habitantes.	Propiedad	LUDLAW
Tipo Norma	Describe un tipo de norma en el contexto normativo de la educación superior.	Propiedad	LUDLAW
Jerarquía Norma	El tipo de jerarquía de las normas según la Pirámide de Kelsen: Nivel 1: Constitución Nivel 2: Convenios internacionales Nivel 3: Legislación Ordinaria: Códigos en General Nivel 4: Legislación Secundaria: Decretos, reglamentos, estatutos, resoluciones	Propiedad	LUDLAW
Lenguaje	Define el idioma de la norma.	Propiedad	DC
Presidencia	Nombre del presidente que en su administración fue creada las normas de educación superior.	Propiedad	FOAF
Creador	La entidad encargada principalmente de crear la norma.	Propiedad	DC
Nombre Registro	Nombre asignado al registro que contiene y publica las normas.	Propiedad	LUDLAW
Numero Registro	Numero asignado al registro oficial.	Propiedad	LUDLAW

Url Norma	Es la dirección online donde se encuentra almacenada la norma.	Propiedad	LUDLAW
Nombre País	Es el nombre de cada país.	Propiedad	LUDLAW
Uri País	Es el uri Dbpedia del país.	Propiedad	LUDLAW
Establece	Las normas establecen a los organismos de educación Superior.	Propiedad	LUDLAW
Nombre Organismo	Nombre del organismo.	Propiedad	FOAF
Descripción Organismo	Una reseña del organismo.	Propiedad	DC
Siglas Organismo	Abreviatura del nombre del organismo.	Propiedad	LUDLAW
Fecha Fundación	Es la fecha en que el organismo fue fundado.	Propiedad	LUDLAW
Homepage	Una página web para algo.	Propiedad	FOAF
Tipo Organismo	Clasificación de los organismos.	Propiedad	LUDLAW
Logo Organismo	Elemento gráfico que identifica a un organismo.	Propiedad	LUDLAW
Número Artículo	Parte numerada de un documento oficial.	Propiedad	LUDLAW
Texto Artículo	Es el texto que define a un artículo.	Propiedad	LUDLAW
Categoría	Es una síntesis de lo que contiene el artículo.	Propiedad	DC
Tag	Palabras que describen un tópico.	Propiedad	DC
Nombre Universidad	Nombre de la universidad.	Propiedad	FOAF
Uri Universidad	Es el uri de la Universidad en Dbpedia.	Propiedad	LUDLAW
Compuesta de	La norma está compuesta por artículos.	Propiedad	LUDLAW
Is parte de	Los artículos son parte de una norma.	Propiedad	LUDLAW
Tiene Sistema	Cada país tiene un Sistema de Educación.	Propiedad	LUDLAW

FASE 3

2.4. Limpieza de Datos.

Limpieza de datos o Data Cleansing es el proceso de detectar, descubrir y corregir datos corruptos, incoherentes o erróneos de un conjunto de datos (Curto, 2008). Después del proceso la información será consistente con otros conjuntos similares de datos. Dicha limpieza de datos puede generar un conjunto de datos más pequeño que el original, lo cual tiene como ventaja mejorar la eficiencia del proceso de Minería de Datos; los datos de calidad deben cumplir los siguientes requisitos: exactitud, integridad, validez, consistencia y uniformidad.

De la información recopilada referente al proyecto sobre Linked Data en el ámbito de Legislación Universitaria, se realizó la limpieza de datos recopilados sobre normas de educación superior de algunos países de Latinoamérica; la información recopilada sobre las normas se la encontró en archivos como PDF, hojas de excel, documentos de Word y la información de organismos de acreditación de la calidad, de coordinación y consulta de educación superior se la extrajo de los diferentes sitios oficiales de los organismos. Luego de completar la información en las hojas de Excel, se procedió a realizar la fase de limpieza de datos para lo cual se utilizó la herramienta Google Refine³⁷ por las siguientes características(Alcubierre, 2012):

- Permite realizar limpieza de un volumen de datos mayor.
- Permite realizar limpieza de la información proveniente de diferentes fuentes de datos.
- Se encuentra disponible para varios sistemas operativos: Windows, Linux, Mac OS.
- Admite varios formatos: TSV, CSV, *SV, Excel (.xls y .xlsx), JSON, XML, RDF como XML, y documentos Google Data.
- Permite transformar los datos a varios formatos.

³⁷<https://github.com/OpenRefine/OpenRefine/wiki/Downloads>

Para realizar la limpieza de los datos debemos identificar cual es la columna de datos a hacer procesada, luego de ello hacemos clic en la pestaña y se desplegarán algunas opciones y se debe escoger Facet y luego Text Facet como la ilustración 20:

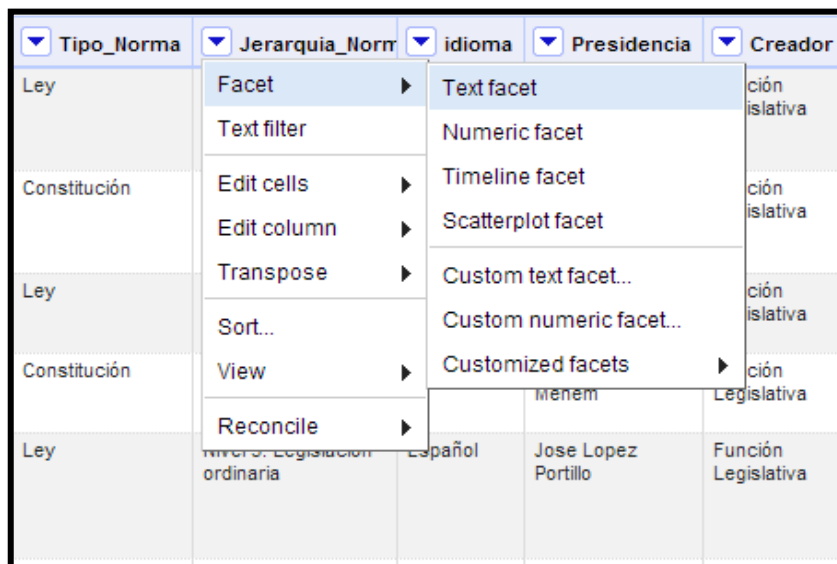


Ilustración 20 Opciones de Google Refine

Luego nos presentará las siguientes pantallas con los datos inconsistentes:

- Varios nombres para el mismo tipo de organismos de educación.

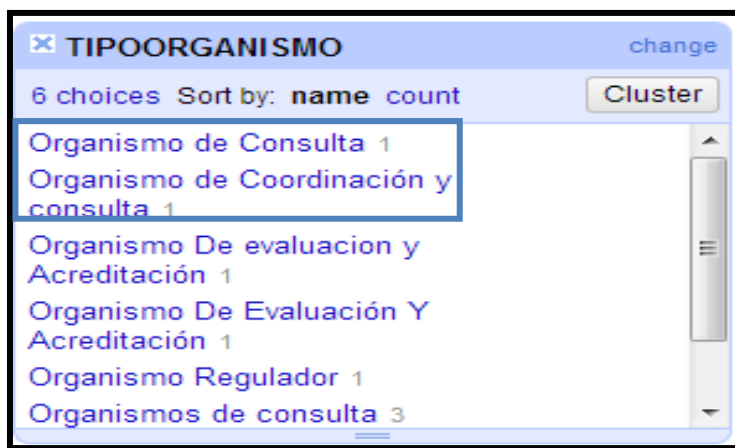


Ilustración 21 Datos Inconsistentes sobre Organismos

- El mismo nombre pero diferentes tipos de escritura, sobre la jerarquía de la norma

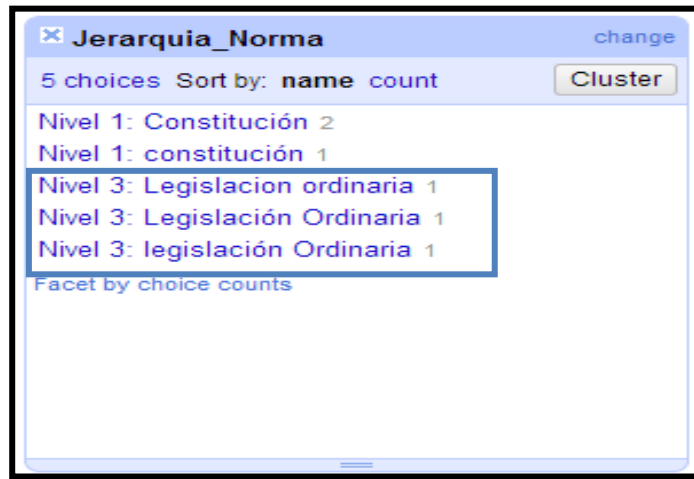


Ilustración 22 Datos Inconsistentes sobre Normas

Para solucionar este inconveniente se lo puede realizar de dos maneras:

- Forma manual: se selecciona lo que se desea cambiar y se edita este cambio se lo realiza uno por uno.
- Forma automática: se da clic en el **botón Cluster** y se presenta la siguiente pantalla en la cual se escoge el método y la función mediante la cual se realizará el cambio automático de los datos inconsistentes.

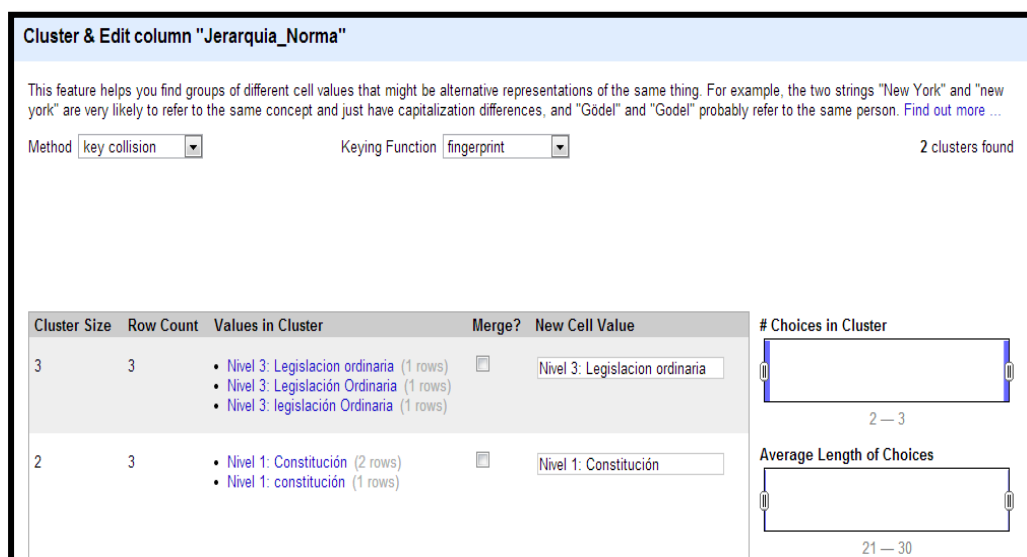


Ilustración 23 Forma Automática mediante Cluster

Luego de realizar este proceso se cuenta con datos consistentes lo cual genera mayor confianza en la información.

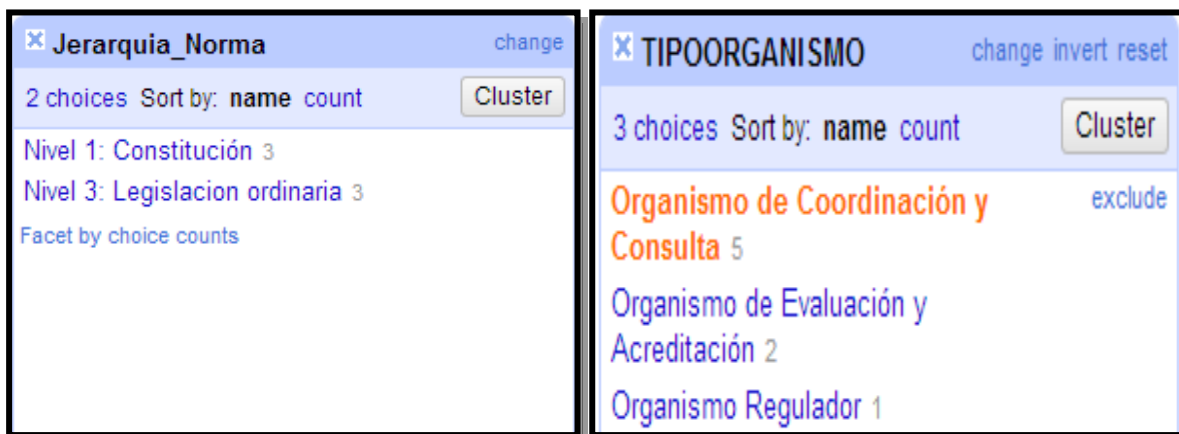


Ilustración 24 Datos Consistentes

Debido a que los datos fueron recolectados de múltiples fuentes es necesario realizar la limpieza de datos con lo cual se contará con datos consistentes; para realizar este proceso se utilizó la herramienta Google Refine ya que permite que el proceso de limpieza y corrección de datos sea eficiente.

Luego de realizar la limpieza de datos se los almacenará en la base de datos MYSQL denominada Leyes los cuales se los convertirá a formato RDF en la siguiente fase de generación de datos.

FASE 4

2.5. Generación de datos a RDF y RDFS.

Siguiendo el proceso de publicación de datos el siguiente paso es la generación de datos a RDF (Corcho & Gómez, 2010).

2.5.1. Generación del RDFS.

El Schema RDF permite definir y describir las clases y propiedades que pueden utilizarse dentro del dominio de Legislación de Educación Superior.

Para el vocabulario LUDLAW se utilizó las siguientes propiedades de RDFS (Tabla 3): `rdfs:Class`, `rdfs:Resource`, `rdfs:label`, `rdfs:comment`, `rdfs:domain`, `rdfs:range` y `rdfs:Datatype`. El RDFS se encuentra en formato de N-Triples (Ver Anexo 4).

- En la primera parte se define los prefijos con los cuales se va a trabajar entre ellos están: RDF, RDFS, XSD, DC, FOAF.
- Luego se definen las clases con cada una de sus propiedades a ser representadas.
- Y también se define los rangos y dominios.

2.5.2. Generación de los datos a RDF.

Para la generación de datos a formato RDF se utilizará Jena, el api de Java que permite crear, manipular grafos RDF y se puede representar recursos, propiedades y literales. Este api permite crear aplicaciones sobre Web Semántica.

Jena tiene integrados algunos de los vocabularios a reutilizar como son: FOAF, DUBLIN CORE, RDF, RDFS debido a que estos vocabularios tienen propiedades que permiten representar la información de Legislación de Educación Superior y para los términos faltantes se creó un nuevo vocabulario denominado LUDLAW.

2.5.2.1. Esquema de la aplicación.

La aplicación que se realizó para convertir los datos a formato RDF, esta creada en la plataforma Java, mediante el API de Jena; para realizar la conversión de los datos se debe acceder a través de una conexión MYSQL (Ver Anexo 2), a la base denominada Leyes, estos datos fueron recolectados de diferentes lugares de la Web, como se puede observar en la ilustración 25.

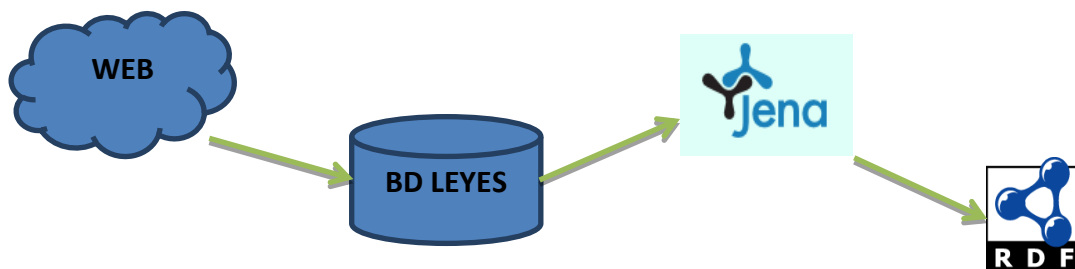
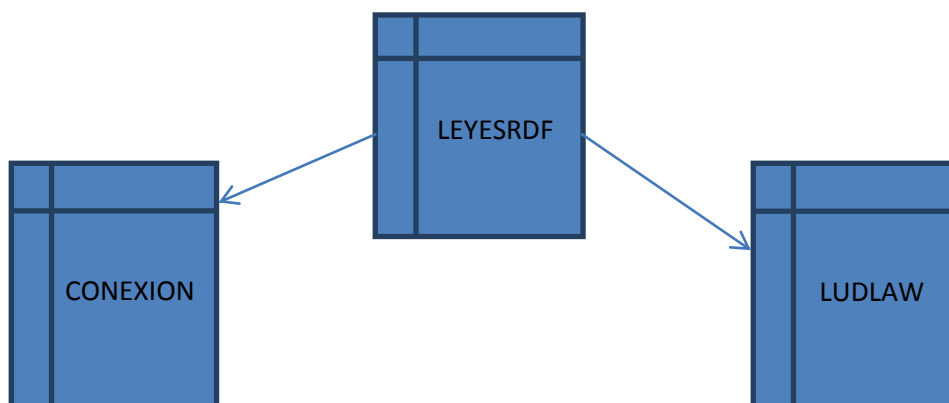


Ilustración 25 Esquema de la aplicación

La aplicación contiene las siguientes clases:



En la clase conexión se conecta Java con la base de datos MYSQL donde están almacenados los datos de Legislación de Educación Superior a ser convertidos a RDF.

El programa contiene un paquete denominado **com.educacion.leyes.vocabulary** el cual tiene una clase denominada LUDLAW en donde se definen las clases y propiedades faltantes para la construcción del vocabulario como se puede observar en la ilustración 26:

```

public class LUDLAW {

    private static Model model = ModelFactory.createDefaultModel();

    public static final String NS = "http://utpl.edu.ec/educacion/leyes/ludlaw/";

    private static String getURI(){
        return NS;
    }

    //Se define las clases del vocabulario
    public static final Resource Norm = model.createResource(NS + "Norm" );
    public static final Resource Article = model.createResource(NS + "Article");
    public static final Resource Country = model.createResource(NS + "Country");
    public static final Resource University = model.createResource(NS + "University");
    public static final Resource SistemaEduSup = model.createResource(NS + "SistemaEduSup");
    public static final Resource Organism = model.createResource(NS + "Organism");
    //Se define las propiedades del vocabulario
    public static final Property promulgationDate = model.createProperty(NS, "promulgationDate");
    public static final Property publicationDate = model.createProperty(NS, "publicationDate");
    public static final Property normType = model.createProperty(NS, "normType");
    public static final Property normHierarchy = model.createProperty(NS, "normHierarchy");
}

```

Ilustración 26 Código del Vocabulario LUDLAW

En la clase LUDLAW para definir los recursos o clases del vocabulario se utiliza la palabra Resource y para las propiedades se utiliza Property.

Y en la clase LEYESRDF (Ver Anexo 3) se procede a instanciar y agregar a cada recurso cada una de las propiedades definidas en el vocabulario LUDLAW y también de los existentes en JENA como son FOAF, DUBLIN CORE.

Como resultado se obtiene un archivo .RDF (Ver Anexo 5), el cual contiene los datos de Legislación de Educación Superior en formato estructurado o RDF listos para ser almacenados en Virtuoso.

La fase de generación de datos a RDF es fundamental en el proceso de publicación de datos debido a que en esta fase se obtendrá datos estructurados, los mismos que serán publicados en el RDF STORE: Virtuoso y mediante consultas SPARQL se facilitará dar respuesta a las interrogantes planteadas.

FASE 5

2.6. Publicación de datos RDF.

Luego de haber convertido los datos sobre Legislación de Educación Superior a formato RDF, se procede almacenar en un RDF Store en este caso se utilizará Virtuoso (Ver Anexo 6), ya que es una herramienta de almacenamiento potente y además cuenta con un ENDPOINT SPARQL para poder explotar la información almacenada.

La información de las tripletas sobre Legislación de Educación Superior se encuentra publicada en el servidor Apolo en la siguiente dirección: <http://apolo.utpl.edu.ec:8890/>.

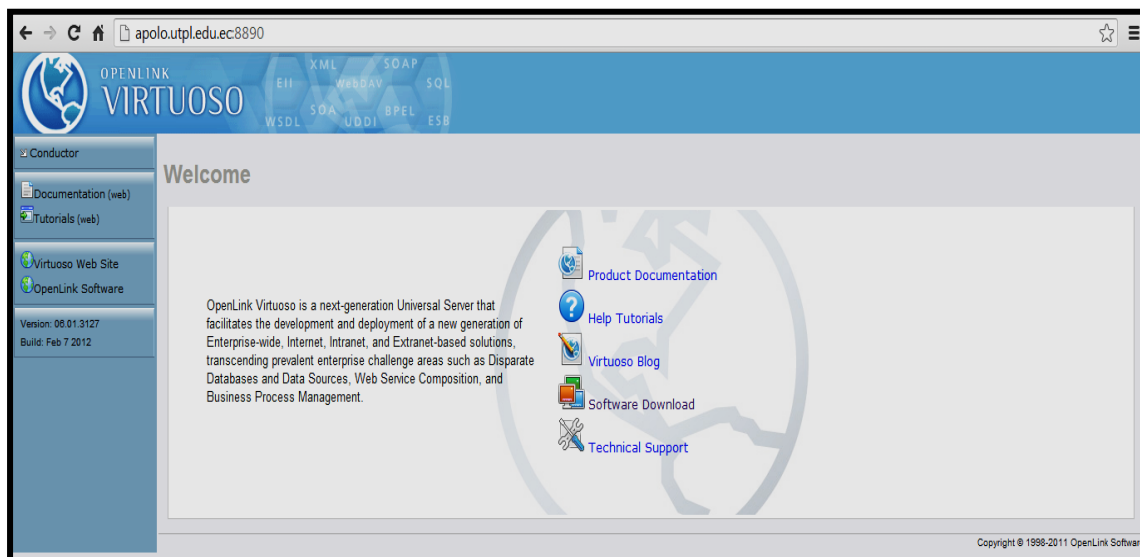


Ilustración 27 Servidor Apolo

2.6.1. Consultas SPARQL.

Luego de haber cargado el archivo RDF (Ver Anexo 7) sobre Legislación de Educación Superior en el motor de almacenamiento de tripletas Virtuoso, accedemos al ENDPOINT SPARQL para lo cual colocamos en el navegador lo siguiente: <http://apolo.utpl.edu.ec:8890/sparql>, previamente debe estar inicializado el servicio de Virtuoso Manager.

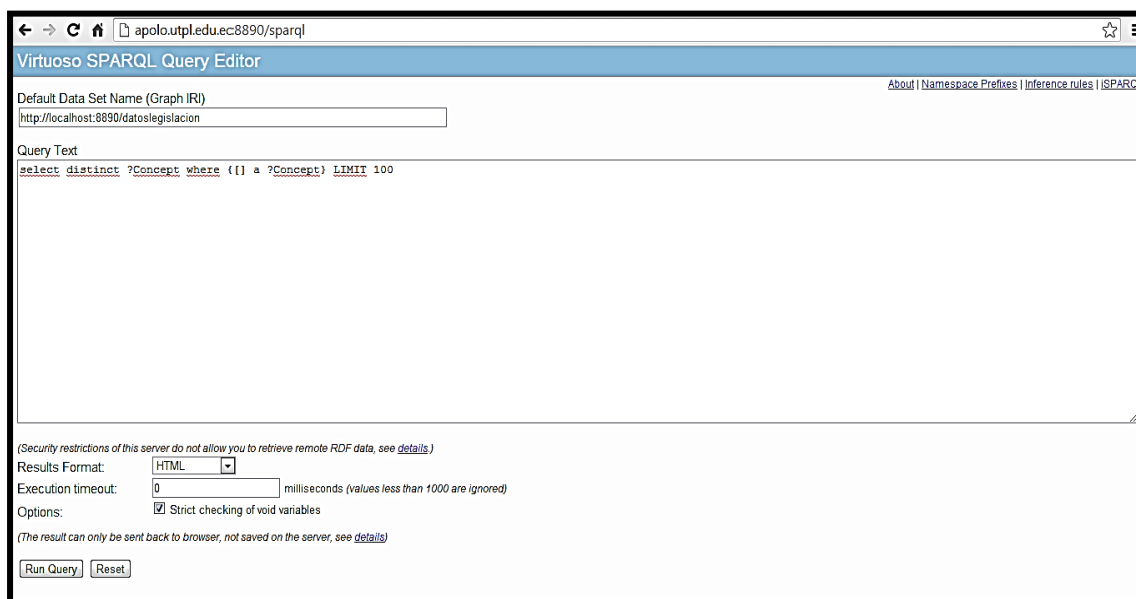


Ilustración 28 EndPoint SPARQL

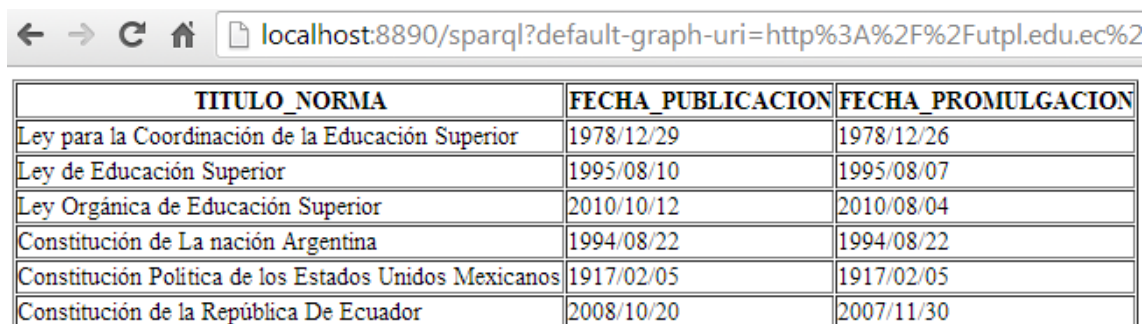
En la ilustración 28 podemos observar el ENDPOINT SPARQL donde se realizará las consultas para dar respuesta a las preguntas planteadas en la parte de especificación de los requisitos de la ontología.

Para realizar las consultas debemos ubicar el IRI del grafo, el cual fue creado al momento de subir el archivo RDF como se puede observar en la ilustración 43. Y en la parte del Query Text ubicamos la consulta SPARQL luego damos clic en Run Query y devolverá el resultadoesperado.

Mediante las respuestas a las interrogantes planteadas al inicio de la investigación se comprobó que se cumplió lo propuesto en la parte de especificación de requisitos.

¿Cuáles son las normas de educación superior y su fecha de publicación?

```
SELECT ?TITULO_NORMA ?FECHA_PUBLICACION ?FECHA_PROMULGACION
WHERE
{
  ?NORMAS dc:title ?TITULO_NORMA.
  ?NORMAS ludlaw:publicationDate ?FECHA_PUBLICACION.
  ?NORMAS ludlaw:promulgationDate ?FECHA_PROMULGACION
}
```



TITULO_NORMA	FECHA_PUBLICACION	FECHA_PROMULGACION
Ley para la Coordinación de la Educación Superior	1978/12/29	1978/12/26
Ley de Educación Superior	1995/08/10	1995/08/07
Ley Orgánica de Educación Superior	2010/10/12	2010/08/04
Constitución de La nación Argentina	1994/08/22	1994/08/22
Constitución Política de los Estados Unidos Mexicanos	1917/02/05	1917/02/05
Constitución de la República De Ecuador	2008/10/20	2007/11/30

Ilustración 29 Resultado Normas

¿Cuáles son las normas que en el articulado mencionado sobre derecho de educación?

```
SELECT ?TITULONORMA ?NUMERO_ARTICULO ?TEXTO_ARTICULO ?CATEGORIA
WHERE
{
  ?ARTICULO ludlaw:isPartOf ?NORMA.
  ?NORMA dc:title ?TITULONORMA
  FILTER REGEX(?ARTICULO, "http://utpl.edu.ec/educacion/leyes/ludlaw/Article/")
  ?ARTICULO ludlaw:isPartOf ?NORMA
  FILTER REGEX(?NORMA, "http://utpl.edu.ec/educacion/leyes/ludlaw/Norm/")
  ?ARTICULO ludlaw:textArticle ?TEXTO_ARTICULO.
  ?ARTICULO ludlaw:numberArticle ?NUMERO_ARTICULO.
  ?ARTICULO dc:description ?CATEGORIA
  FILTER REGEX(?CATEGORIA, "DERECHO EDUCACION")
}
```

TITULONORMA	NUMERO_ARTICULO	TEXTO_ARTICULO	CATEGORIA
Ley de Educación Superior	Art. 11	Son derechos de los docentes estatales de las instituciones estatales de educación superior, sin perjuicio de lo dispuesto por la legislación específica: a) Acceder a la carrera académica mediante concurso público y abierto de antecedentes y oposición; b) Participar en el gobierno de la institución a la que pertenecen, de acuerdo a las normas legales pertinentes; c) Actualizarse y perfeccionarse de modo continuo a través de la carrera académica; d) Participar en la actividad gremial.	DERECHO EDUCACION
Ley de Educación Superior	Art. 13	Los estudiantes de las instituciones estatales de educación superior tienen derecho: a) Al acceso al sistema sin discriminaciones de ninguna naturaleza. b) A asociarse libremente en centros de estudiantes, federales nacionales y regionales, a elegir sus representantes y a participar en el gobierno y en la vida de la institución, conforme a los estatutos, lo que establece la presente ley y, en su caso, las normas legales de las respectivas jurisdicciones; c) A obtener becas, créditos y otras formas de apoyo económico y social que garanticen la igualdad de oportunidades y posibilidades, particularmente para el acceso y permanencia en los estudios de grado, conforme a las normas que reglamenten la materia; d) A recibir información para el adecuado uso de la oferta de servicio de educación superior; e) A solicitar, cuando se encuentren en las situaciones previstas en los artículos 1ero y 2do de la ley 20.596, la postergación o adelanto de exámenes o evaluaciones parciales o finales cuando las fechas previstas para los mismos se encuentren dentro del periodo de preparación y/o participación.	DERECHO EDUCACION
Ley Orgánica de Educación Superior	Art. 4	Derecho a la Educación Superior.- El derecho a la educación superior consiste en el ejercicio efectivo de la igualdad de oportunidades, en función de los méritos respectivos, a fin de acceder a una formación académica y profesional con producción de conocimiento pertinente y de excelencia. Las ciudadanas y los ciudadanos en forma individual y colectiva, las comunidades, pueblos y nacionalidades tienen el derecho y la responsabilidad de participar en el proceso educativo superior, a través de los mecanismos establecidos en la Constitución y esta Ley.	DERECHO EDUCACION
Ley Orgánica de Educación Superior	Art. 5	Derechos de las y los estudiantes.- Son derechos de las y los estudiantes los siguientes: a) Acceder, movilizarse, permanecer, egresar y titularse sin discriminación conforme sus méritos académicos; b) Acceder a una educación superior de calidad y pertinente, que permita iniciar una carrera académica y/o profesional en igualdad de oportunidades; c) Contar y acceder a los medios y recursos adecuados para su formación superior; garantizados por la Constitución; d) Participar en el proceso de evaluación y acreditación de su carrera; e) Elegir y ser elegido para las representaciones estudiantiles e integrar el cogobierno, en el caso de las universidades y escuelas politécnicas; f) Ejercer la libertad de asociarse, expresarse y completar su formación bajo la más amplia libertad de cátedra e investigativa; g) Participar en el proceso de construcción, difusión y aplicación del conocimiento; h) El derecho a recibir una educación superior laica, intercultural, democrática, incluyente y diversa, que impulse la equidad de género, la justicia y la paz; e, i) Obtener de acuerdo con sus méritos académicos becas, créditos y otras formas de apoyo económico que le garantice igualdad de oportunidades en el proceso de formación de educación superior.	DERECHO EDUCACION

Ilustración 30 Resultado Normas sobre Derecho Educación

¿Cuáles son los países que en su normativa mencionan sobre el derecho de la Autonomía Universitaria?

```

SELECT ?PAIS ?NORMAS ?NUMERO_ARTICULO ?TEXTO_ARTICULO ?CATEGORIA
WHERE
{
?ARTICULO ludlaw:isPartOf ?NORMA.
?NORMA dc:title ?NORMAS
    FILTER REGEX(?ARTICULO, "http://utpl.edu.ec/educacion/leyes/ludlaw/Article/")
    FILTER REGEX(?NORMA, "http://utpl.edu.ec/educacion/leyes/ludlaw/Norm/")
?ARTICULO dc:description ?CATEGORIA
    FILTER REGEX(?CATEGORIA, "AUTONOMIA UNIVERSITARIA")
    FILTER REGEX(?ARTICULO, "http://utpl.edu.ec/educacion/leyes/ludlaw/Article/")
?SISTEMAEDU ludlaw:regulatedBy ?NORMA.
    FILTER REGEX(?SISTEMAEDU,
"http://utpl.edu.ec/educacion/leyes/ludlaw/SistemaEducacion/")
?PAISES ludlaw:hasSystem ?SISTEMAEDU.
?ARTICULO ludlaw:textArticle ?TEXTO_ARTICULO.
?ARTICULO ludlaw:numberArticle ?NUMERO_ARTICULO.
?PAISES ludlaw:nameCountry ?PAIS
}

```

PAIS	NORMAS	NUMERO ARTICULO	TEXTO ARTICULO	CATEGORIA
Ecuador	Ley Orgánica de Educación Superior	Art. 18	Ejercicio de la autonomía responsable.- La autonomía responsable que ejercen las universidades y escuelas politécnicas consiste en: a) La independencia para que los profesores e investigadores de las universidades y escuelas politécnicas ejerzan la libertad de cátedra e investigación; b) La libertad de expedir sus estatutos en el marco de las disposiciones de la presente Ley; c) La libertad en la elaboración de sus planes y programas de estudio en el marco de las disposiciones de la presente Ley; d) La libertad para nombrar a sus autoridades, profesores o profesoras, investigadores o investigadoras, las y los servidores y las y los trabajadores, atendiendo a la alternancia y equidad de género, de conformidad con la Ley; e) La libertad para gestionar sus procesos internos; f) La libertad para elaborar, aprobar y ejecutar el presupuesto institucional. Para el efecto, en el caso de instituciones públicas, se observarán los parámetros establecidos por la normativa del sector público; g) La libertad para adquirir y administrar su patrimonio en la forma prevista por la Ley; h) La libertad para administrar los recursos acorde con los objetivos del régimen de desarrollo, sin perjuicio de la fiscalización a la institución por un órgano contralor interno o externo, según lo establezca la Ley; e, i) La capacidad para determinar sus formas y órganos de gobierno, en consonancia con los principios de alternancia, equidad de género, transparencia y derechos políticos señalados por la Constitución de la República, e integrar tales órganos en representación de la comunidad universitaria, de acuerdo a esta Ley y los estatutos de cada institución.	AUTONOMIA UNIVERSITARIA
Ecuador	Ley Orgánica de Educación Superior	Art. 19	Inviolabilidad de los recintos universitarios.- Los recintos de las universidades y escuelas politécnicas son inviolables y no podrán ser allanados sino en los casos y términos en que puede serlo el domicilio de una persona, según lo previsto en la Constitución y la Ley. Deben servir exclusivamente, para el cumplimiento de sus fines y objetivos definidos en esta Ley. La vigilancia y el mantenimiento del orden interno son de competencia y responsabilidad de sus autoridades. Cuando se necesite el resguardo de la fuerza pública, el representante legal de la institución solicitará la asistencia pertinente, de lo cual informará en su momento al órgano colegiado académico superior. Quienes violaren estos recintos serán sancionados de conformidad con la Ley.	AUTONOMIA UNIVERSITARIA
Ecuador	Ley Orgánica de Educación Superior	Art. 17	Reconocimiento de la autonomía responsable.- El Estado reconoce a las universidades y escuelas politécnicas autonomía académica, administrativa, financiera y orgánica, acorde con los principios establecidos en la Constitución de la República. En el ejercicio de autonomía responsable, las universidades y escuelas politécnicas mantendrán relaciones de reciprocidad y cooperación entre ellas y de estas con el Estado y la sociedad; además observarán los principios de justicia, equidad, solidaridad, participación ciudadana, responsabilidad social y rendición de cuentas.	AUTONOMIA UNIVERSITARIA
Ecuador	Constitución de la República De Ecuador	Art. 355	El Estado reconocerá a las universidades y escuelas politécnicas autonomía académica, administrativa, financiera y orgánica, acorde con los objetivos del régimen de desarrollo y los principios establecidos en la Constitución. Se reconoce a las universidades y escuelas politécnicas el derecho a la autonomía, ejercida y comprendida de manera solidaria y responsable. Dicha autonomía garantiza el ejercicio de la libertad académica y el derecho a la búsqueda de la verdad, sin restricciones; el gobierno y gestión de sí mismas, en consonancia con los principios de alternancia, transparencia y los derechos políticos; y la producción de ciencia, tecnología, cultura y arte. Sus recintos son inviolables, no podrán ser allanados sino en los casos y términos en que pueda serlo el domicilio de una persona. La garantía del orden interno será competencia y responsabilidad de sus autoridades. Cuando se necesite el resguardo de la fuerza pública, la máxima autoridad de la entidad solicitará la asistencia pertinente. La autonomía no exime a las instituciones del sistema de ser fiscalizadas, de la responsabilidad social, rendición de cuentas y participación en la planificación nacional. La Función Ejecutiva no podrá privar de sus rentas o asignaciones presupuestarias, o retardar las transferencias a ninguna institución del sistema, ni clausurarlas o reorganizarse de forma total o parcial.	AUTONOMIA UNIVERSITARIA
Argentina	Ley de Educación Superior	Art. 32	Contra las resoluciones definitivas de las instituciones universitarias nacionales, impugnadas con fundamento en la interpretación de las leyes de la Nación, los estatutos y demás normas internas, sólo podrá interponerse recurso de apelación ante la Cámara Federal de Apelaciones con competencia en el lugar donde tiene su sede principal la institución universitaria.	AUTONOMIA UNIVERSITARIA
Argentina	Ley de Educación Superior	Art. 31	La fuerza pública no puede ingresar en las instituciones universitarias nacionales si no mediante orden escrita previa y fundada de juez competente o solicitud expresa de la autoridad universitaria legítimamente constituida.	AUTONOMIA UNIVERSITARIA

Ilustración 31 Resultado Países sobre derecho Autonomía Universitaria

¿Cuáles son las universidades que son parte del sistema de educación superior de Ecuador?

```
SELECT ?PAIS ?NOMBRE_UNIVERSIDAD ?URI_UNIVERSIDAD
WHERE
{
?UNIVERSIDAD ludlaw:isPartOf ?SISTEMAEDU
FILTER REGEX(?UNIVERSIDAD,
"http://utpl.edu.ec/educacion/leyes/ludlaw/University/")
?UNIVERSIDAD foaf:name ?NOMBRE_UNIVERSIDAD.
?UNIVERSIDAD ludlaw:uriUniversity ?URI_UNIVERSIDAD.
?PAISES ludlaw:hasSystem ?SISTEMAEDU.
?PAISES ludlaw:nameCountry ?PAIS
FILTER REGEX(?PAIS, "Ecuador")
}
```

PAIS	NOMBRE UNIVERSIDAD	URI UNIVERSIDAD
Ecuador	Universidad T% <u>C3%A9</u> cnica Estatal de Quevedo	http://dbpedia.org/resource/Universidad_T% <u>C3%A9</u> cnica Estatal de Quevedo
Ecuador	Universidad T% <u>C3%A9</u> cnica Particular de Loja	http://dbpedia.org/resource/Universidad_T% <u>C3%A9</u> cnica Particular de Loja
Ecuador	Escuela Superior Politecnica del Litoral	http://dbpedia.org/resource/Escuela_Superior_Politecnica_del_Litoral
Ecuador	Escuela Polit% <u>C3%A9</u> cnica del Ej% <u>C3%A9</u> rcito	http://dbpedia.org/resource/Escuela_Polit% <u>C3%A9</u> cnica del Ej% <u>C3%A9</u> rcito
Ecuador	University of Cuenca	http://dbpedia.org/resource/University_of_Cuenca
Ecuador	National Polytechnic School	http://dbpedia.org/resource/National_Polytechnic_School
Ecuador	Central University of Ecuador	http://dbpedia.org/resource/Central_University_of_Ecuador
Ecuador	University of Guayaquil	http://dbpedia.org/resource/University_of_Guayaquil
Ecuador	Politecnica Salesiana University	http://dbpedia.org/resource/Politecnica_Salesiana_University
Ecuador	UEES	http://dbpedia.org/resource/UEES
Ecuador	Universidad Del Pac% <u>C3%AD</u> fico %E2% <u>80%93</u> Ecuador	http://dbpedia.org/resource/Universidad_Del_Pac% <u>C3%AD</u> fico_%E2% <u>80%93</u> Ecuador
Ecuador	Universidad San Francisco de Quito	http://dbpedia.org/resource/Universidad_San_Francisco_de_Quito

Ilustración 32 Resultado Universidades del Sistema Educación

¿En qué normas se definen los organismos de educación superior?

```

SELECT ?PAIS ?TITULO_NORMA ?NUMERO_ARTICULO ?TEXTO_ARTICULO ?CATEGORIA
WHERE
{
  ?ARTICULO ludlaw:isPartOf ?NORMA.
  ?NORMA dc:title ?TITULO_NORMA
    FILTER REGEX(?ARTICULO, "http://utpl.edu.ec/educacion/leyes/ludlaw/Article/")
    FILTER REGEX(?NORMA, "http://utpl.edu.ec/educacion/leyes/ludlaw/Norm/")
  ?ARTICULO ludlaw:numberArticle ?NUMERO_ARTICULO.
  ?ARTICULO ludlaw:textArticle ?TEXTO_ARTICULO.
  ?ARTICULO dc:description ?CATEGORIA
    FILTER REGEX(?CATEGORIA, "ORGANISMOS DE EDUCACION")?SISTEMAEDU
  ludlaw:regulatedBy ?NORMA.
  FILTER REGEX(?SISTEMAEDU,
"http://utpl.edu.ec/educacion/leyes/ludlaw/SistemaEducacion/")
  ?PAISES ludlaw:hasSystem ?SISTEMAEDU.
  ?PAISES ludlaw:nameCountry ?PAIS }

```

PAIS	TITULO_NORMA	NUMERO_ARTICULO	TEXTO_ARTICULO	CATEGORIA
Ecuador	Ley Orgánica de Educación Superior	Art. 194	Los Comités Regionales Consultivos de Planificación de la Educación Superior serán órganos de consulta regional de la Secretaría Nacional de Educación Superior, Ciencia, Tecnología e Innovación, de articulación con el trabajo desconcentrado de la Función Ejecutiva y de coordinación territorial con los actores de la educación superior que trabajen a escala regional y de los gobiernos regionales autónomos. Su finalidad es constituirse en herramienta de consulta horizontal del Sistema de Educación Superior a nivel regional, para hacer efectiva la articulación territorial con el resto de niveles y modalidades educativas del Sistema Educativo Nacional y las distintas áreas gubernamentales de necesaria interacción con las instituciones de nivel superior, tales como la planificación nacional y regional, la ciencia, la tecnología y la producción. Funcionará un Comité Regional Consultivo de Planificación de la Educación Superior por cada región autónoma que se constituya en el país.	ORGANISMOS DE EDUCACION
Ecuador	Ley Orgánica de Educación Superior	Art. 171	Consejo de Evaluación, Acreditación y Aseguramiento de la Calidad de la Educación Superior. - El Consejo de Evaluación, Acreditación y Aseguramiento de la Calidad de la Educación Superior es el organismo público técnico, con personería jurídica y patrimonio propio, con independencia administrativa, financiera y operativa. Funcionará en coordinación con el Consejo de Educación Superior. Tendrá facultad regulatoria y de gestión. No podrá conformarse por representantes de las instituciones objeto de regulación ni por aquellos que tengan intereses en las áreas que vayan a ser reguladas.	ORGANISMOS DE EDUCACION
Ecuador	Ley Orgánica de Educación Superior	Art. 166	Consejo de Educación Superior. - El Consejo de Educación Superior es el organismo de derecho público con personería jurídica, con patrimonio propio, independencia administrativa, financiera y operativa, que tiene por objetivo la planificación, regulación y coordinación interna del Sistema de Educación Superior, y la relación entre sus distintos actores con la Función Ejecutiva y la sociedad ecuatoriana. El Consejo de Educación Superior funcionará en coordinación con el Consejo de Evaluación, Acreditación y Aseguramiento de la Calidad de la Educación Superior y no podrá conformarse por representantes de las instituciones objeto de regulación ni por aquellos que tengan intereses en las áreas que vayan a ser reguladas. Tendrá su sede en la capital de la República.	ORGANISMOS DE EDUCACION
Ecuador	Ley Orgánica de Educación Superior	Art. 185	Asamblea del Sistema de Educación Superior. - La Asamblea del Sistema de Educación Superior es el órgano representativo y consultivo que sugiere al Consejo de Educación Superior, políticas y lineamientos para las instituciones que conforman el Sistema de Educación Superior. Con fines informativos, conocerá los resultados de la gestión anual del Consejo.	ORGANISMOS DE EDUCACION
Argentina	Ley de Educación Superior	Art. 73	El Consejo Interuniversitario Nacional estará integrado por los rectores o presidentes de las instituciones universitarias nacionales y provinciales reconocidas por la Nación, que estén definitivamente organizadas, y el Consejo de Rectores de Universidades Privadas estará integrado por los rectores o presidentes de las instituciones universitarias privadas. Dichos consejos tendrán por funciones: a) Coordinar los planes y actividades en materia académica, de investigación científica y de extensión entre las instituciones universitarias de sus respectivos ámbitos; b) Ser órganos de consulta en las materias y cuestiones que prevé la presente ley; c) Participar en el Consejo de Universidades. Cada Consejo se dará su propio reglamento conforme al cual regulará su funcionamiento interno.	ORGANISMOS DE EDUCACION
Argentina	Ley de Educación Superior	Art. 46	La Comisión Nacional de Evaluación y Acreditación Universitaria es un organismo descentralizado, que funciona en jurisdicción del Ministerio de Educación, y que tiene por funciones: a) Coordinar y llevar adelante la evaluación externa prevista en el artículo 44; b) Acreditar las carreras de grado a que se refiere el artículo 43, así como las carreras de posgrado, cualquiera sea el ámbito en que se desarrollen, conforme a los estándares que establece el Ministerio de Educación en consulta con el Consejo de Universidades; c) Pronunciarse sobre la consistencia y viabilidad del proyecto institucional que se requiere para el Ministerio de Educación autorice la puesta en marcha de una nueva institución universitaria nacional con la posterioridad a su creación o el reconocimiento de una institución universitaria provincial; d) Preparar los informes requeridos para otorgar la autorización provisoria y el reconocimiento definitivo de las instituciones universitarias privadas, así como los informes en base a los cuales se evaluará el periodo de funcionamiento provisorio de dichas instituciones.	ORGANISMOS DE EDUCACION
Argentina	Ley de Educación Superior	Art. 70	Corresponde al Ministerio de Educación la formulación de las políticas generales en materia universitaria, asegurando la participación de los órganos de coordinación y consulta previsto en la presente ley y respetando el régimen de autonomía establecido para las instituciones universitarias.	ORGANISMOS DE EDUCACION
México	Ley para la Coordinación de la Educación Superior	Art. 15	Habrà un Consejo del Sistema Nacional de Educación Tecnológica que será órgano de consulta de la Secretaría de Educación Pública, de las entidades federativas cuando éstas lo soliciten y de las instituciones públicas de educación tecnológica de tipo superior, para coordinar las actividades de dicho sistema y contribuir a vincularlas con las necesidades y el desarrollo del país.	ORGANISMOS DE EDUCACION
México	Ley para la Coordinación de la Educación Superior	Art. 14	Habrà un Consejo Nacional Consultivo de Educación Normal, cuya integración determinará el Ejecutivo Federal, que será órgano de consulta de la Secretaría de Educación Pública, de las entidades federativas cuando éstas lo soliciten y de las instituciones de educación normal para coordinar sus actividades, orientar la celebración de los convenios que sobre la materia prevé esta ley y contribuir a vincular dicha educación con los requerimientos del país, de conformidad con la política educativa nacional.	ORGANISMOS DE EDUCACION

Ilustración 33 Resultado Normas que definen Organismos

En el Anexo 8 pueden observar otras consultas SPARQL sobre el ámbito de Legislación de Educación Superior.

FASE 6

2.7. Enlazado de datos.

Luego de realizar la conversión de datos a formato RDF se procede a realizar el enlazado para lo cual se puede seguir el siguiente proceso (Gómez & Corcho, 2012):

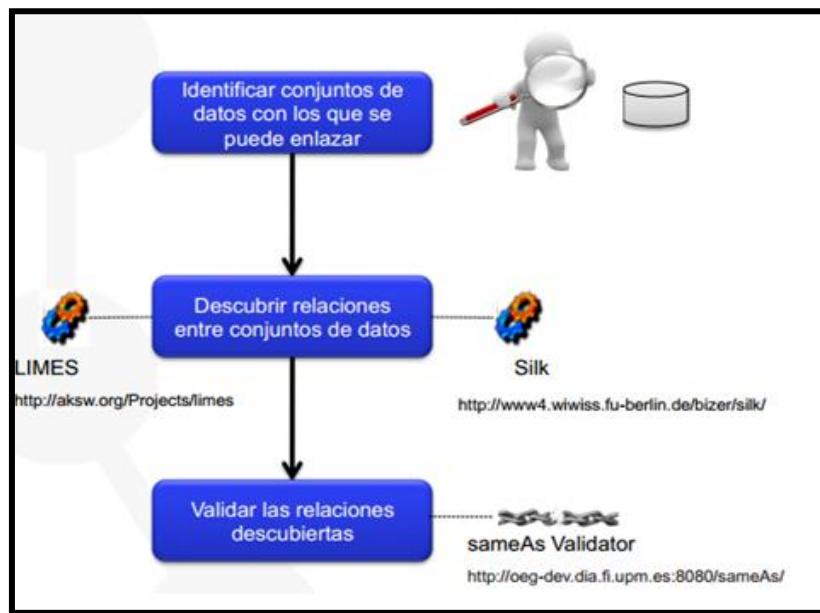


Ilustración 34 Proceso de enlazado de datos (Gómez & Corcho, 2012)

Como se puede observar en la ilustración 34:

El primer paso del proceso de enlazado de datos consiste en identificar los dataset que se encuentre en diferentes RDF Store y disponibles para enlazar.

El segundo paso consiste en descubrir las relaciones entre conjuntos de datos para lo cual se puede utilizar la propiedad de owl:SameAs que se utiliza para unir dos instancias que se refieren a lo mismo.

Para realizar el enlazado de información o Linking se pueden utilizar algunas herramientas genéricas tales como: SILK³⁸ esta herramienta permite establecer enlaces RDF de sus fuentes de datos con otras fuentes de datos disponibles en la Web, LAMES³⁹ es un framework para enlace de datos.

³⁸<http://wifo5-03.informatik.uni-mannheim.de/bizer/silk/>

³⁹<http://aksw.org/Projects/LIMES.html>

La información que se intenta enlazar es la siguiente:

Organismos de educación superior

Se realizó la búsqueda de información sobre Organismos de Educación Superior en los más grandes RDF Store como son: DBpedia, Freebase, Sindice la búsqueda se la realizó por el nombre y por las siglas de los organismos de ninguna forma se obtuvo ningún resultado, los organismos que se buscaron son los siguientes:

Tabla 11 Búsqueda de organismos en Datasets

ORGANISMOS	DBPEDIA	FREEBASE	SINDICE
Comisión Nacional de Evaluación y Acreditación Universitaria (CONEAU)	X	X	X
Consejo de Evaluación, Acreditación y Aseguramiento de la Calidad de la Educación Superior(CEAACES)	X	X	X
Consejo de Educación Superior(CES)	X	X	X
Consejo Nacional de Ciencia y Tecnología (CONACYT)	X	X	X
Secretaría Nacional de Ciencia y Tecnología (SENESCYT)	X	X	X

La información sobre países se encuentra disponible en diferentes datasets y puede ser enlazada utilizando la propiedad owl:sameAs, con los datasets que se pueden observar en la ilustración 35:

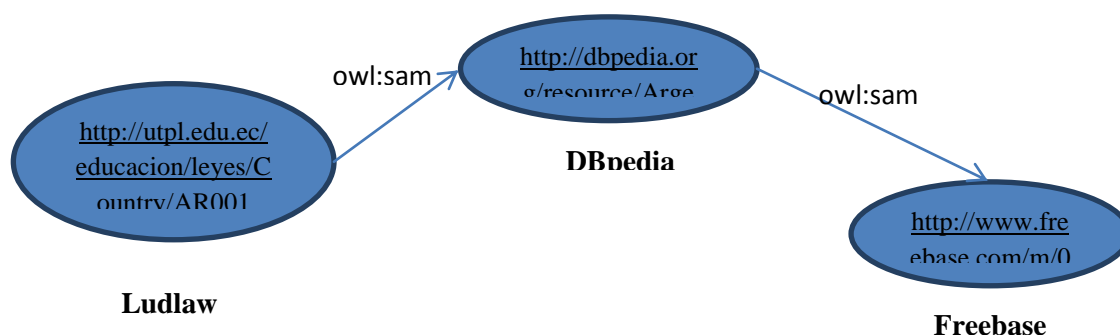


Ilustración 35 Datasets sobre Países

En la tabla 12 se encuentra los uris de los países hacer enlazados, tienen diferente escritura pero representa al mismo país.

Tabla 12 Uris de los países

País	DBPEDIA	FREEBASE	LUDLAW
Argentina	http://dbpedia.org/resource/Argentina	http://www.freebase.com/m/0jgd	http://utpl.edu.ec/educacion/leyes/Country/AR001
Ecuador	http://dbpedia.org/resource/Ecuador	http://www.freebase.com/m/02k1b	http://utpl.edu.ec/educacion/leyes/Country/EC001
Mexico	http://dbpedia.org/resource/Mexico	http://www.freebase.com/m/0b90_r	http://utpl.edu.ec/educacion/leyes/Country/MX001

Las clases que se utilizaron para el enlazado de países son:

- **dbpedia-owl:Country**
- **ludlaw:Country**

Para realizar el enlazado debemos ingresar la información del Dataset con el cual se va enlazar los datos como puede observar en la ilustración 36:

Ilustración 36 Source para Enlazar

Luego de haber ingresado el endpointURI nos aparecerá la siguiente pantalla la cual indica algunos parámetros que fueron configurados:

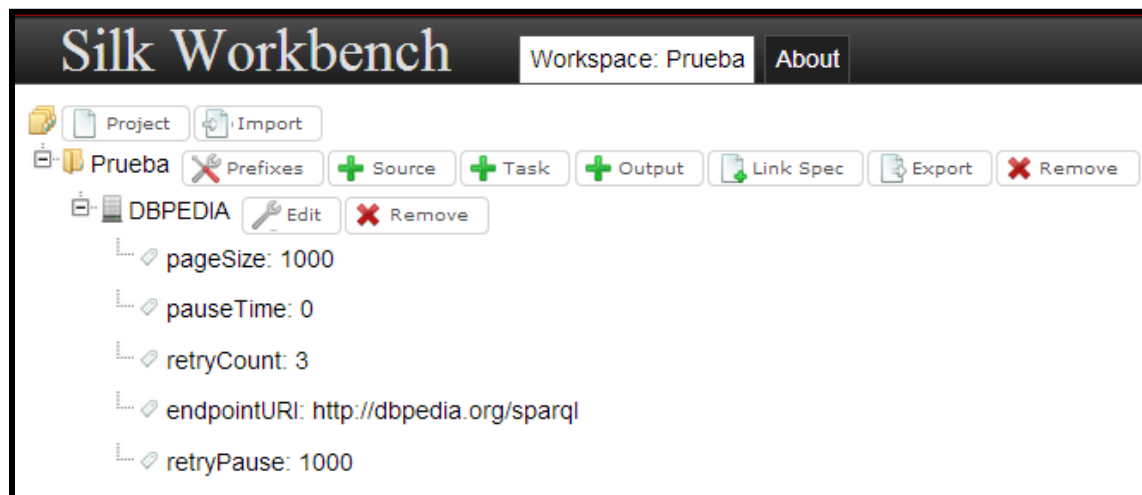


Ilustración 37 Dataset de DBpedia

Después de agregar los datasets con los que se va a enlazar se debe crear una tarea de Linking en la cual se debe seleccionar el origen(Source) y destino(target) de los datos a enlazar y para cada una se debe realizar una consulta SPARQL y mediante la propiedad owl:sameAs se realizará el enlazado.

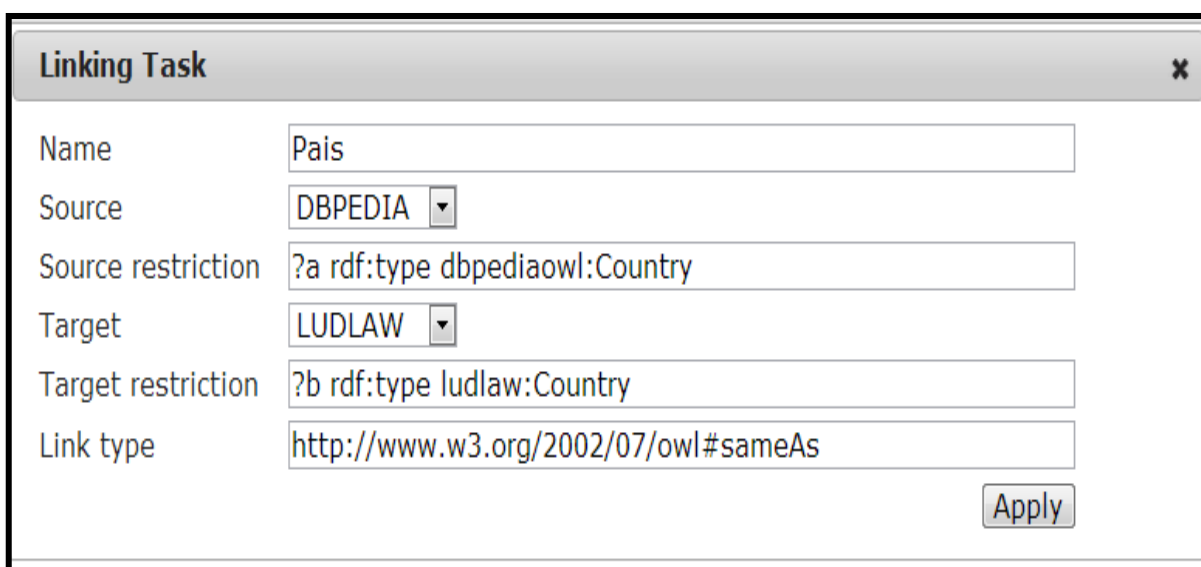


Ilustración 38 Enlace de datos

Otra información que se encontró en la web para enlazar es sobre los nombres de los presidentes que en su administración crearon las normas de educación superior y se enlazará con los datos publicados en Dbpedia y freebase⁴⁰, como se puede observar en la ilustración 39.

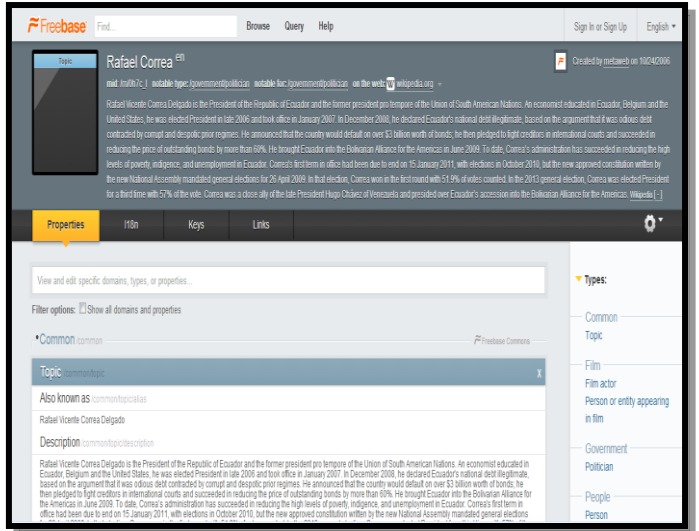
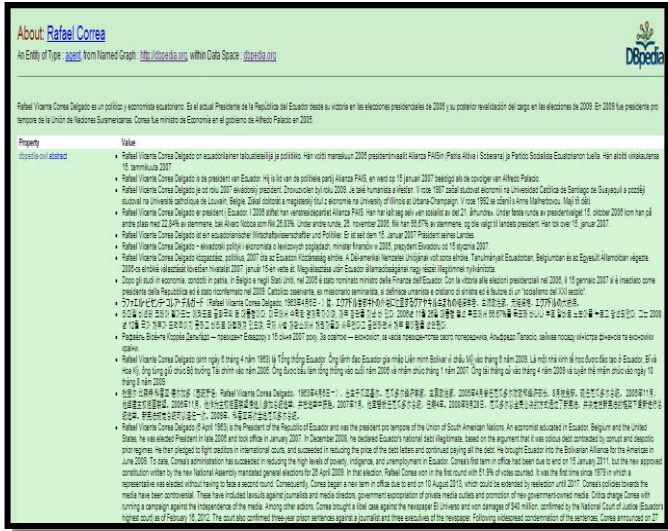


Ilustración 39 DBpedia y Freebase

⁴⁰ <http://www.freebase.com>

DISCUSIÓN

En el presente proyecto de investigación se realizó la Publicación de datos enlazados en el ámbito de Legislación de Educación Superior de los países: Ecuador, Argentina y México de los cuales se extrajo información de los tipos de norma: Ley y Constitución utilizando los principios de Linked Open Data de tal manera que estos puedan ser accedidos fácilmente; los datos se encuentran almacenados en Virtuoso, para lograr este objetivo se utilizará algunas de las tecnologías semánticas mencionadas en el estado del arte.

La parte fundamental de la discusión se centra en la fase de recolección de datos debido a que la información se encuentra suelta, redundante y de calidad dudosa además la interoperabilidad por formatos dificulta que la extracción sea de manera automática; la recolección de forma manual implica mucho tiempo y retraso en las demás fases de la publicación de datos enlazados, para evitar este inconveniente se puede utilizar técnicas de análisis de textos como es el Procesamiento de Lenguaje Natural (PLN), mediante esta técnica se puede disminuir el tiempo empleado en la extracción y además se contaría con un formato semántico según la normativa de Linked Data.

La recolección de datos de los Organismos de Educación Superior se lo realizó manualmente debido a que eran pocos sitios web los que se analizaron; si el tamaño de organismos aumentará es aconsejable utilizar la técnica de Web Scraping para recolectar datos de la web de forma automática con el fin de agilizar el proceso y cubrir las necesidades.

En lo relacionado a la fase del desarrollo de vocabulario se puede concluir que la reutilización de vocabularios existentes permite que los datos ya no estén aislados dentro de un contexto único.

Mediante las consultas SPARQL se busca agilizar y encontrar solución a las interrogantes planteadas al inicio de la investigación; con la finalidad de comprobar se realizó una prueba haciendo la búsqueda en Google sobre ¿Cuáles son las normas que en su articulado mencionan sobre el derecho de la Autonomía Universitaria? y arrojó 158.000 resultados de los cuales el usuario tendría que ir revisando y elegir cual es el resultado correcto; esto conllevaría mucho tiempo y resulta fastidioso para el usuario; mientras que al realizar la

búsqueda en Virtuoso nos presente datos concretos ya que estos datos se encuentran en un formato estructurado y el tiempo de búsqueda es más corto.

A futuro en este proyecto se puede obtener la información de manera automática sobre Legislación de Educación Superior de todos los países de Latinoamérica ya que en esta versión solo se extrajo de tres países Ecuador, Argentina y Mexico; también se puede agregar otros tipos de normas como Decretos, Reglamentos, Resolución, Estatutos y Acuerdos con la finalidad de contar con un dataset completo y estructurado sobre Legislación Universitaria.

CONCLUSIONES

Después de haber realizado el proyecto de tesis sobre publicación de datos enlazados en el ámbito de Legislación de Educación Superior se concluye lo siguiente:

- Mediante la aplicación de los principios de publicación de datos enlazados, estándares como lenguaje de consulta (SPARQL) y lenguaje de representación de datos (RDF), permite realizar una búsqueda exacta de datos referentes a Legislación de Educación Superior debido a que se cuenta con datos en formato estructurado dotados de mayor significado. Lo que se logró en este proyecto fue convertir los datos a RDF, obtener uris y publicar en un SPARQL ENDPOINT los datos de Legislación Universitaria para lo cual se utilizó Virtuoso; mediante el cual se ejecutan consultas SPARQL y ayuda a precisar los resultados de la búsqueda.
- En la fase de identificación y extracción de datos sobre Legislación de Educación Superior para realizar la identificación de datos se requirió de guía de expertos del dominio para cumplir esta fase y se dificulta la extracción de la información puesto que está disponible en documentos no estructurados para lo cual se requiere un proceso más complejo debido a que se encuentra en diferentes formatos digitales y físicos.
- Luego de realizar el análisis de vocabularios se determina que se puede reutilizar los siguientes: Dublin Core, Aiiiso, FOAF y la Ontología de Normas de la Biblioteca del Congreso Nacional de Chile debido a que contiene clases y propiedades que permiten cubrir parte del dominio del proyecto para las propiedades faltantes se construirá el vocabulario denominado LUDLAW; mediante este proceso se logra cumplir uno de los objetivos primordiales de Linked Data que es la reutilización de vocabularios.

- La fase de limpieza de datos es fundamental ya que permitió detectar y corregir datos incoherentes en el ámbito de legislación universitaria lo cual permite contar con datos consistentes e íntegros y de esta manera brindar respuestas acertadas del ámbito de estudio; para realizar la limpieza existen varias herramientas que permiten agilizar el proceso de limpieza el que se utilizó en este proyecto es Google Refine.
- Para la construcción de la ontología se utilizó la metodología NEON por la flexibilidad de adaptación a las necesidades de los usuarios. Los escenarios que se utilizaron en el proyecto son el escenario 1 que se basa en la especificación de requisitos y el escenario 2 que se refiere a la reutilización de recursos no ontológicos como bases de datos por ejemplo de DBPEDIA se extrajo información de la universidades y países de Latinoamérica y vocabularios controlados mediante esta reutilización se trata de reducir el tiempo y coste del desarrollo de ontologías.

RECOMENDACIONES

- Para la publicación de datos enlazados se recomienda utilizar un proceso de publicación de datos el cual incluye los siguientes pasos como son: Identificar y seleccionar fuentes de datos, desarrollo de vocabularios, limpieza de datos, generación de datos a RDF, publicación y enlazado de datos RDF con otras fuentes con la finalidad de contar con datos estructurados y completos que faciliten la búsqueda de la información en el ámbito de Legislación Universitaria.
- Es aconsejable realizar un análisis previo de los proyectos relacionados en el ámbito de Legislación Universitaria, con el objetivo de que si existen trabajos relacionados, vocabularios u ontologías se los reutilice con el fin de agilizar el proceso de construcción del vocabulario.
- Para obtener los requisitos necesarios para la construcción del vocabulario es necesario contar con un experto en el dominio sobre Legislación Universitaria y así obtener un modelo que se asemeje a la realidad y satisfacer las necesidades propuestas.

TRABAJOS FUTUROS

- Desarrollar un software de análisis y extracción de datos sobre Legislación Universitaria; de manera que la extracción se la pueda realizar mediante criterios especificados.
- Utilizar la técnica de Web Scraping para obtener datos de las páginas web sobre los organismos de educación superior.
- Desarrollar una aplicación que permita visualizar los datos sobre Legislación Universitaria a través de las consultas SPARQL con la finalidad de que los usuarios potenciales puedan acceder fácilmente a la información.

GLOSARIO DE TÉRMINOS

- **RDF:** lenguaje estándar para representar el conocimiento humano en formato digital.
- **URI:** identificador único de recursos, es el conjunto genérico de todos los nombres y direcciones que se refieren a un recurso.
- **JAVA:** es una plataforma virtual de software de tal manera que los programas creados en ella puedan ejecutarse sin cambios en diferentes tipos de arquitecturas. Es la primera plataforma informática creada por Sun Microsystems en 1995.
- **SPARQL:** es lenguaje de consultas para RDF.
- **MYSQL:** es un sistema de gestión de bases de datos relacional, multihilo y multiusuario.
- **RDFS (Esquema RDF):** definición de propiedades y relaciones entre recursos identificados en la Word WideWeb.
- **JENA:** es un API para Java, es decir, un entorno para el desarrollo de aplicaciones Java para la web semántica.
- **Metadato:** es un dato que se utiliza para describir o añadir información sobre otro dato.
- **Dublin Core:** es un estándar para representar metadatos sobre recursos web.
- **N3:** tipo de formato de RDF.
- **Namespace (Espacio de Nombres):** permite identificar una fuente que define un conjunto de elementos y atributos que son utilizados en documento XML.
- **Tripleta:** es una estructura que contiene sujeto, predicado y objeto.
- **Property:** es un atributo de un recurso. Por ejemplo DC.language es una propiedad, como es RDF.type.
- **API:** interfaz de programación de aplicaciones.
- **HTTP:** es el protocolo de transferencia de hipertexto, que se usa en cada transacción en la web.

BIBLIOGRAFÍA

Akoma Ntoso. (2011). XML FOR Parliamentary, legislative and judiciary documents. Recuperado el 10 de 04 de 2013, de <http://www.akomantoso.org/>

Alcubierre, D. (2012). Google Refine, más allá del excel. Recuperado el 15 de 06 de 2013, de <http://www.theproject.ws/en/toolab/post/google-refine-mas-alla-del-excel>

Barrera, M., & Núñez, H. (01 de 2012). Ingeniería Ontológica. Recuperado el 23 de 05 de 2013

Benjamins, R., & Casanovas, P. (12 de 06 de 2012). Law and the Semantic Web, an Introduction. Obtenido de <http://alumnos.dcc.uchile.cl/~ddiaz/Papers/Otros/fulltext.pdf>

Berners-Lee. (2007). Hearing on the "Digital Future of the United States: Part I. Recuperado el 26 de 11 de 2012, de <http://dig.csail.mit.edu/2007/03/01-ushouse-future-of-the-web.pdf>

Berners-Lee. (25 de 03 de 2012). The World Wide Web: Past, Present and Future. Obtenido de <http://www.w3.org/People/Berners-Lee/1996/ppf.html>

Berners-Lee, T. (2005). Uniform resource identifier (URI): generic syntax. Recuperado el 12 de 02 de 2013, de <http://www.ietf.org/rfc/rfc3986.txt>

Berners-Lee, T. (04 de 13 de 2012). Linked data. Obtenido de <http://www.w3.org/DesignIssues/LinkedData.html>

Biblioteca del Congreso Nacional de Chile. (12 de 08 de 2011). Ontología de Normas para la Biblioteca del Congreso Nacional de Chile. Recuperado el 17 de 03 de 2013, de <http://datos.bcn.cl/ontologies/bcn-norms/doc/images/ontologia.gif>

Bizer, C. (25 de 04 de 2012). How to publish linked data on the web. Obtenido de <http://www4.wiwiw.fu-berlin.de/bizer/pub/LinkedDataTutorial>

Bizer, C., Cyganiak, R., & Heath, T. (2007). How to Publish Linked Data on the Web. Recuperado el 04 de 09 de 2013, de <http://wifo5-03.informatik.uni-mannheim.de/bizer/pub/LinkedDataTutorial/>

Bizer, C., Heath, T., & Berners, L. (03 de 2009). Linked Data - The Story So Far. Recuperado el 13 de 01 de 2013, de <http://tomheath.com/papers/bizer-heath-berners-lee-ijswis-linked-data.pdf>

Casellas, N. (15 de 03 de 2010). Semantic Enhancement of Legal Information: Are We Up for the Challenge? Obtenido de <http://blog.law.cornell.edu/voxpath/2010/02/15/semantic-enhancement-of-legal-information%e2%80%a6-are-we-up-for-the-challenge>

Casellas, N. (09 de 05 de 2011). A Case Study in U.S. Federal Legal Information.

Casellas, N. (2012). Legally Linked: Linked Open Data Principles Applied To Code Of Federal Regulations. Obtenido de http://semanticweb.com/legally-linked-linked-open-data-principles-applied-to-code-of-federal-regulations_b29197

Casellas, N., Vallbé, J., & Bruce, T. (2011). From Legal Information to Open Legal Data. Recuperado el 25 de 08 de 2012, de <http://dx.doi.org/10.2139/ssrn.1959931>

Castells, P., & Camacho, D. (2006). La web semántica: tecnologías y aplicaciones. Recuperado el 26 de 12 de 2012, de http://www.kramirez.net/SMA_Maestria/Material/Presentaciones/0-websemantica.pdf

CEN MetaLex. (2010). Open XML Interchange Format for Legal and Legislative Resources. Recuperado el 20 de 06 de 2013, de <http://www.metalex.eu/>

Cifuentes, F. (12 de 08 de 2011). Ontología de Normas para la Biblioteca del Congreso Nacional de Chile.

Cifuentes, F. (2011). Publicación de Datos Abiertos Enlazados En El Ámbito Legislativo. Recuperado el 24 de 06 de 2012, de http://156.35.31.156/~fcifuentes/fileadmin/user_upload/tesis_master.pdf

Classora Technologies. (05 de 11 de 2012). Sacando provecho a la Web Semántica: SPARQL. Recuperado el 22 de 03 de 2013, de <http://blog.classora.com/2012/11/05/sacando-provecho-a-la-web-semantica-sparql/>

Corcho, O., & Gómez, A. (2010). Mini-curso sobre Linked Data. Recuperado el 25 de 02 de 2012, de <http://www.slideshare.net/ocorcho/linked-data-tutorial-florianopolis>

Cornell University Law School. (2010). Legal Information Institute Open Acces to Law Since 1992. Recuperado el 19 de 04 de 2013, de <http://www.law.cornell.edu/>

DCMI. (01 de 05 de 2012). Expressing Simple Dublin Core in RDF/XML. Recuperado el 26 de 03 de 2013, de <http://dublincore.org/documents/dcmes-xml/>

DuCharme, B. (2011). Learning SPARQL. Beijing: O'Reilly Media, Inc.

Estrella. (2008). Estrella project website. Recuperado el 25 de 06 de 2013, de <http://www.estrellaproject.org/>

Fernandez, M., Gomez, A., & Juristo, N. (1997). METHONTOLOGY: From Ontological Art Towards Ontological Engineering. American Asociation for Artificial Intelligence.

Franz Incorporated. (2011). AllegroGraph 4.11 Introduction. Recuperado el 01 de 03 de 2013, de <http://www.franz.com/agraph/support/documentation/current/agraph-introduction.html>

Gómez, A., & Corcho, O. (13 de 12 de 2012). Herramientas de publicación de datos para portales de datos abiertos. Recuperado el 24 de 08 de 2013, de http://datos.gob.es/datos/sites/default/files/files/%2004.V3_Tecnologia_Portales_OpenData.pdf

ITTIG. (2010). Institute of Legal Information Theory and Techniques. Recuperado el 08 de 04 de 2013, de <http://www.ittig.cnr.it/Ricerca/UnitaEng.php?Id=40>

Jena. (2011). Apache Jena. Recuperado el 16 de 03 de 2013, de <http://jena.apache.org/>

Leibniz Center for Law. (2010). LRICore Ontology. Recuperado el 26 de 04 de 2013, de <http://www.leibnizcenter.org/previous-projects/lricore>

Méndez, E. (24 de 04 de 2012). "RDF: un modelo de metadatos flexible para las bibliotecas digitales del próximo milenio". Recuperado el 14 de 03 de 2013, de <http://www.cobdc.org/jornades/7JCD/1.pdf>

Neon. (24 de 02 de 2010). Welcome to the NeOn Project. Recuperado el 30 de 06 de 2013, de http://www.neon-project.org/nw/Welcome_to_the_NeOn_Project

OpenLink Software. (2010). OpenLink Virtuoso Universal Server: Documentation. Recuperado el 22 de 02 de 2013, de <http://docs.openlinksw.com/pdf/virtdocs.pdf>

Oreilly, T. (22 de 03 de 2012). What is web 2.0: Design patterns and business models for the next generation of software. Obtenido de <http://oreilly.com/web2/archive/what-is-web-20.html>

Rob Style, & Nadeem Shabir. (2008). Academic Institution Internal Structure Ontology (AIISO). Recuperado el 24 de 08 de 2012, de <http://vocab.org/aiiso/schema>

Rubio, A. (2006). Historia E Internet: Aproximación Al Futuro De La Labor Investigativa. Obtenido de <http://www.ucm.es/info/hcs/angel/articulos/historiaeinternet.pdf>

Salazar, J. (2011). Estado actual de la Web 3.0 o Web Semántica. Revista Digital Universitaria, 12(11), 17.

Sesame. (2010). System documentation for Sesame 2. Recuperado el 17 de 02 de 2013, de <http://openrdf.callimachus.net/sesame/2.7/docs/system.docbook?view>

Styles, R., & Shabir, N. (2008). Academic Institution Internal Structure Ontology (AIISO). Recuperado el 28 de 03 de 2013, de <http://vocab.org/aiiso/schema>

Suárez, M. (2010). Tesis Doctoral: NeOn Methodology for Building Ontology Networks: Specification, Scheduling and Reuse. Recuperado el 01 de 07 de 2013

Suárez-Figueroa. (2008). NeOn methodology for Building. Recuperado el 24 de 06 de 2013, de <http://www.neon-project.org>

Tim Berners, L., & Hendler, J. (26 de 04 de 2001). "Publishing on the semantic Web" Nature. Obtenido de <http://www.nature.com/nature/journal/v410/n6832/full/4101023b0.html>

W3C. (02 de 10 de 2004). RDF Primer. Recuperado el 14 de 03 de 2013, de <http://www.w3.org/TR/2004/REC-rdf-primer-20040210/>

W3C. (2008). RDF Vocabulary Description Language 1.0: RDF Schema. Recuperado el 20 de 03 de 2013, de <http://www.w3.org/TR/rdf-schema/>

W3C. (28 de 08 de 2012). Linked Data. Obtenido de <http://www.w3.org/standards/semanticweb/data>

Wyner, A. (13 de 06 de 2011). Applying Human Language Technology to the Law. Obtenido de <http://wyner.info/research/Papers/AHLTL2011Papers.pdf>

ANEXOS

Anexo 1. Script de la base leyes desde Power Designer

```
/*=====*/
/* DBMS name:      MySQL 5.0                               */
/* Created on:     11/03/2013 20:47:17                     */
/*=====*/
drop table if exists ARTICULO;
drop table if exists NORMA;
drop table if exists ORGANISMO;
drop table if exists PAIS;
drop table if exists UNIVERSIDADES;
drop table if exist SISTEMA EDUCACION;
/*=====*/
/* Table: ARTICULO                                         */
/*=====*/
create table ARTICULO
(
IDARTICULO          varchar(100) not null,
  IDNORMA            varchar(100),
  NUM_ART            varchar(1024),
TEXTOARTICULO       varchar(5000),
  CATEGORIA          varchar(1024),
  TAGS               varchar(1024),
primary key (IDARTICULO)
);
/*=====*/
/* Table: NORMA                                           */
/*=====*/
create table NORMA
(
IDNORMA             varchar(100) not null,
  IDPAIS             varchar(100),
  TITULO             varchar(1024),
  FECHA_PROMULGACION varchar(1024),
  FECHA_PUBLICACION  varchar(1024),
  TIPO_NORMA         varchar(1024),
  JERARQUIA_NORMA    varchar(1024),
  IDIOMA             varchar(1024),
PRESIDENCIA         varchar(1024),
CREADOR             varchar(1024),
  LUGAR_PUBLI        varchar(1024),
  NUMERO             varchar(1024),
  URL                varchar(1024),
```

```

primary key (IDNORMA)
);
/*=====*/
/* Table: ORGANISMO */
/*=====*/
create table ORGANISMO
(
    IDORGANISMO      varchar(100) not null,
    IDNORMA          varchar(100),
    NOMBRE_ORG      varchar(1024),
    DESCRIPCION     varchar(5000),
    TIPO_ORG        varchar(1024),
    SIGLAS          varchar(1024),
    FUNDACION       varchar(1024),
    LOGO            varchar(1024),
    HOMEPAGE        varchar(1024),
primary key (IDORGANISMO)
);
/*=====*/
/* Table: PAIS */
/*=====*/
create table PAIS
(
    IDPAIS          varchar(100) not null,
    PAIS            varchar(100),
    URI_PAIS       Varchar(1024),
primary key (IDPAIS)
);
/*=====*/
/* Table: UNIVERSIDADES */
/*=====*/
create table UNIVERSIDADES
    IDUNI          varchar(100) not null,
    IDPAIS        varchar(100),
    URIUNIVERSIDAD  varchar(1024),
    NOMBREUNIVER  varchar(1024),
primary key (IDUNI)
);
/*=====*/
/* Table: SISTEMA EDUCACION */
/*=====*/
create table SISTEMA EDUCACION
    IDSIS        varchar(250) not null,

```

```
IDPAIS    varchar(50),
SISTEMAEDUC    varchar(1024),

primary key (IDSIS)
);
alter table ARTICULO add constraint FK_CONTIENE foreign key (IDNORMA)
references NORMA (IDNORMA) on delete restrict on update restrict;
alter table NORMA add constraint FK_PAIS_TIENE_NORMA foreign key (IDPAIS)
references PAIS (IDPAIS) on delete restrict on update restrict;
alter table ORGANISMO add constraint FK_DEFINE foreign key (IDNORMA)
references NORMA (IDNORMA) on delete restrict on update restrict;
alter table SISTEMA EDUCACION add constraint FK_PAIS_SISTEMA foreign key (IDPAIS)
references PAIS (IDPAIS) on delete restrict on update restrict;
alter table UNIVERSIDADES add constraint FK_PAIS_UNIVERSIDADES foreign key (IDPAIS)
references PAIS (IDPAIS) on delete restrict on update restrict;
```

Anexo 2. Conexión MYSQL desde JENA

```
package conexion;
import java.io.*;
import java.sql.*;
import com.mysql.jdbc.exceptions.MySQLDataException;
/**
 *
 * @author MAFER
 */
public class Conexion {
    public Connection coneccion ;
    private static Conexion objeto = null;
    private Conexion() throws Exception
    {
        Conexion();
    }
    public static Conexion Singleton() throws Exception
    {
        if (objeto == null)
            objeto = new Conexion();
        return objeto;
    }
    private void Conexion() throws Exception
    {
        BufferedReader bu;
        int cont=0;
        String aux=null;
        String driver = "com.mysql.jdbc.Driver";
        String user = "root";
        String password="";
        String url = "";
        try{
            bu = new BufferedReader(new FileReader("C:\\systemdb.txt"));
            while ((aux=bu.readLine())!=null)
            {
                if(cont==0)
                    user=aux;
                if(cont==1)
                    password=aux;
                if(cont==2)
                    url=aux;
                cont++;
            }
            Class.forName(driver).newInstance();
            coneccion =DriverManager.getConnection(url,user,password);
        }
        catch(MySQLDataException ex){
            throw ex;
        }
    }
}
```

Anexo 3. Código para generación de RDF

```
package datosrdf;
import com.hp.hpl.jena.rdf.model.Model;
import com.hp.hpl.jena.rdf.model.ModelFactory;
import com.hp.hpl.jena.rdf.model.Resource;
import com.hp.hpl.jena.sparql.vocabulary.FOAF;
import com.hp.hpl.jena.vocabulary.DC;
import com.hp.hpl.jena.vocabulary.RDF;
import com.hp.hpl.jena.vocabulary.RDFS;
import com.educacion.leyes.vocabulary.LUDLAW;
import com.hp.hpl.jena.vocabulary.XSD;
import com.mysql.jdbc.PreparedStatement;
import conexion.Conexion;
import java.io.FileWriter;
import java.io.IOException;
import java.sql.ResultSet;
import java.text.SimpleDateFormat;
import java.util.Calendar;
/**
 *
 * @author MAFER
 */
public class LeyesRDF {

    public static void main(String[] args) throws Exception {
        boolean existe = false;
        PreparedStatement smt;

        ResultSet rs;
        Model model = ModelFactory.createDefaultModel();
        model.setNsPrefix("foaf", FOAF.NS);
        model.setNsPrefix("ludlaw", LUDLAW.NS);

        String strsql = "SELECT
IDPAIS,PAIS,URI_PAIS,IDSIST,SISTEMAEDUC, IDNORMA,TITULO,FECHA_PROMULGACION,FECHA_PUBLICACION,
TIPO_NORMA,JERARQUIA_NORMA,IDIOMA,PRESIDENCIA,CREADOR,LUGAR_PUBLICACION,NUMERO,URL,
IDARTICULO,NUM_ART,TEXTOARTICULO,CATEGORIA,TAGS, IDORGANISMO,NOMBRE_ORG,DESCRIPCIONORG,
TIPO_ORG,SIGLAS,FUNDACION,LOGO,HOMEPAge,IDUNI,URIUNIVERSIDAD,NOMBREUNIVER FROM
datosleyes";
        Conexion con = Conexion.Singleton();
        java.sql.Statement sql = con.coneccion.createStatement();
        existe = sql.executeQuery(strsql).next();
        rs = sql.executeQuery(strsql);
```

```

if (existe != false) {
    while (rs.next()) {
        String idpais = rs.getString("IDPAIS");
        String pais = rs.getString("PAIS");
        String uri_pais = rs.getString("URI_PAIS");
        String idsis = rs.getString("IDSIST");
        String sistedu = rs.getString("SISTEMAEDUC");
        String idnorma = rs.getString("IDNORMA");
        String titulo = rs.getString("TITULO");
        String fechapromul = rs.getString("FECHA_PROMULGACION");
        String fechapubli = rs.getString("FECHA_PUBLICACION");
        String tipoNorma = rs.getString("TIPO_NORMA");
        String jerarquiaNorma = rs.getString("JERARQUIA_NORMA");
        String idioma = rs.getString("IDIOMA");
        String nombrePresidente = rs.getString("PRESIDENCIA");
        String creador = rs.getString("CREADOR");
        String nombreRegistro = rs.getString("LUGAR_PUBLI");
        String numero = rs.getString("NUMERO");
        String urlNorma = rs.getString("URL");
        String idarticulo = rs.getString("IDARTICULO");
        String numeroart = rs.getString("NUM_ART");
        String textoArt = rs.getString("TEXTOARTICULO");
        String descripcionArt = rs.getString("CATEGORIA");
        String tags = rs.getString("TAGS");
        String idorganismo = rs.getString("IDORGANISMO");
        String nombreorg = rs.getString("NOMBRE_ORG");
        String tipoorg = rs.getString("TIPO_ORG");
        String siglasorg = rs.getString("SIGLAS");
        String descripOrg = rs.getString("DESCRIPCIONORG");
        String fundacionorg = rs.getString("FUNDACION");
        String logourl = rs.getString("LOGO");
        String pagina = rs.getString("HOMEPAGE");
        String iduniv = rs.getString("IDUNI");
        String uriuniv = rs.getString("URIUNIVERSIDAD");
        String nombreuni = rs.getString("NOMBREUNIVER");
        //Se define el recurso Country
        String paisURI = "http://utpl.edu.ec/educacion/leyes/ludlaw/Country/"
+ idpais;

        Resource Country = model.createResource(paisURI)
            .addProperty(LUDLAW.nameCountry, pais)
            .addProperty(LUDLAW.uriCountry, uri_pais);

        //Se define el recurso Sistema de Educacion Superior
        String sistemaURI =
"http://utpl.edu.ec/educacion/leyes/ludlaw/SistemaEducacion/" + idsis;

```



```

Resource SistemaEducacion = model.createResource(sistemaURI)
    .addProperty(LUDLAW.systemEducation, sistedu);
Country.addProperty(LUDLAW.hasSystem, SistemaEducacion);

//Se define el recurso Universidad
String universidadURI =
"http://utpl.edu.ec/educacion/leyes/ludlaw/University/" + iduniv;
Resource University = model.createResource(universidadURI)
    .addProperty(FOAF.name, nombreuni)
    .addProperty(LUDLAW.uriUniversity, uriuniv)
    .addProperty(RDF.type, LUDLAW.University);
University.addProperty(LUDLAW.isPartOf, SistemaEducacion);

//Se define el recurso Norma
String normaURI = "http://utpl.edu.ec/educacion/leyes/ludlaw/Norm/" +
idnorma;

Resource Norm = model.createResource(normaURI)
    .addProperty(DC.title, titulo)
    .addProperty(LUDLAW.promulgationDate, fechapromul)
    .addProperty(LUDLAW.publicationDate, fechapubli)
    .addProperty(LUDLAW.normType, tipoNorma)
    .addProperty(LUDLAW.normHierarchy, jerarquiaNorma)
    .addProperty(FOAF.name, nombrePresidente)
    .addProperty(DC.creator, creador)
    .addProperty(LUDLAW.nameRegister, nombreRegistro)
    .addProperty(LUDLAW.numberRegister, numero)
    .addProperty(LUDLAW.urlNorm, urlNorma)
    .addProperty(RDF.type, LUDLAW.Norm)
    .addProperty(RDFS.domain, LUDLAW.Norm)
    .addProperty(RDFS.range, LUDLAW.Article);
SistemaEducacion.addProperty(LUDLAW.regulatedBy, Norm);

//Se define el recurso Artículo
if (idarticulo != null) {

    String artURI =
"http://utpl.edu.ec/educacion/leyes/ludlaw/Article/" + idarticulo;
    Resource Article = model.createResource(artURI)
        .addProperty(LUDLAW.numberArticle, numeroart)
        .addProperty(LUDLAW.textArticle, textoArt)
        .addProperty(DC.description, descripcionArt)
        .addProperty(DC.subject, tags)
        .addProperty(RDF.type, LUDLAW.Article);
    Article.addProperty(LUDLAW.isPartOf, Norm);
}

```

```

//Se define el recurso Organismos
if (idorganismo != null) {
    String orgURI =
"http://utpl.edu.ec/educacion/leyes/ludlaw/Organism/" + idorganismo;
    Resource Organism = model.createResource(orgURI);

    if (nombreorg != null) {
        Resource node = model.createResource(orgURI)
            .addProperty(FOAF.name, nombreorg);
    }

    if (descripOrg != null) {
        Resource node = model.createResource(orgURI)
            .addProperty(DC.description, descripOrg);
    }

    if (siglasorg != null) {
        Resource node = model.createResource(orgURI)
            .addProperty(LUDLAW.orgAcronym, siglasorg);
    }

    if (fundacionorg != null) {
        Resource node = model.createResource(orgURI)
            .addProperty(LUDLAW.foundingDate, fundacionorg);
    }

    if (pagina != null) {
        Resource node = model.createResource(orgURI)
            .addProperty(FOAF.homepage, pagina);
    }

    if (tipoorg != null) {
        Resource node = model.createResource(orgURI)
            .addProperty(LUDLAW.orgType, tipoorg);
    }

    if (logourl != null) {
        Resource node = model.createResource(orgURI)
            .addProperty(LUDLAW.orgLogo, logourl);
    }

    Organism.addProperty(RDF.type, LUDLAW.Organism);

    Organism.addProperty(LUDLAW.established, Norm);

```

```

        }
    }
    model.write(System.out);
    try {
        Calendar cal = Calendar.getInstance();
        SimpleDateFormat sdf = new SimpleDateFormat("'Date: 'EEE, dd MMM yyyy
HH:mm:ss '-0700 ( MST ) '");
        SimpleDateFormat output = new SimpleDateFormat("yyyy-MM-dd hh-mm");
        String dateStr = output.format(cal.getTime());
        FileWriter fstream = new FileWriter("C:\\rdf\\leyes_".concat(dateStr)
+ ".rdf");
        model.write(fstream);
    } catch (IOException e) {
        System.out.println("Exception caught" + e.getMessage());
    }
}
sql.close();
}
}

```

Anexo 4. Vocabulario en formato N3

```
@prefix rdf:<http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs:<http://www.w3.org/2000/01/rdf-schema#> .
@prefix owl:<http://www.w3.org/2002/07/owl#> .
@prefix xsd:<http://www.w3.org/2001/XMLSchema#> .
@prefix skos:<http://www.w3.org/2004/02/skos/core#> .
@prefix dct:<http://purl.org/dc/terms/> .
@prefix dc:<http://purl.org/dc/elements/1.1/> .
@prefix foaf:<http://xmlns.com/foaf/0.1/> .
@prefix sioc:<http://rdfs.org/sioc/ns#> .
@prefix dbpedia:<http://dbpedia.org/resource/> .
@prefix dbpedia-owl:<http://dbpedia.org/ontology/> .
@prefix fb:<http://rdf.freebase.com/ns/> .
@prefix aiiso:<http://purl.org/vocab/aiiso/schema#> .
@prefix aiiso-roles:<http://purl.org/vocab/aiiso-roles/schema#> .
@prefix schema:<http://schema.org/> .
@prefix ludlaw:<http://utpl.edu.ec/educacion/leyes#> .
@prefix bcnnorms:<http://datos.bcn.cl/ontologies/bcn-norms#> .

Norm
:Norm rdf:type rdfs:class .
:Norm rdfs:label "Norm" .
:Norm rdfs:comment "Contiene información sobre las normas que rigen la educación superior en los diferentes países." .

:Norm owl:equivalentClass bcnnorms:Norm .

Title
:title rdf:type rdf:Property .
:title rdfs:label "Title" .
:title rdfs:comment "El título es un nombre con el que el recurso se conoce formalmente." .
:title owl:equivalentProperty dc:title .
:title rdfs:domain ludlaw:Norm.
:title rdfs:range xsd:string .

Promulgation Date
:promulgationDate rdf:type rdf:Property .
:promulgationDate rdfs:label "Promulgation Date" .
:promulgationDate rdfs:comment "Cuando la norma es creada y aprobada por la función Ejecutiva de cada país." .
:promulgationDate owl:equivalentProperty bcnnorms:promulgationDate .
:promulgationDate rdfs:subpropertyOf dc:date .
:promulgationDate rdfs:domain ludlaw:Norma .
:promulgationDate rdfs:range xsd:date .

Publication Date
:publicationDate rdf:type rdf:Property .
:publicationDate rdfs:label "Publication Date" .
:publicationDate rdfs:comment "Cuando la norma es publicada en el periódico oficial de cada país." .

:publicationDate owl:equivalentProperty bcnnorms:publishDate .
```

```

:publicationDate rdfs:subpropertyOf dc:date .
:publicationDate rdfs:domain ludlaw:Norma .
:publicationDate rdfs:range xsd:date .

Norm Type
:normType rdf:type rdf:Property .
:normType rdfs:label "Norm Type" .
:normType rdfs:comment "El tipo de norma que regula la educación superior puede ser: Ley,
constitución." .
:normType rdfs:domain ludlaw:Norma .
:normType rdfs:range xsd:string .

Norm Hierarchy
:normHierarchy rdf:type rdf:Property .
:normHierarchy rdfs:label "Norm Hierarchy" .
:normHierarchy rdfs:comment "Clasificación de las normas por niveles como: en el nivel
jerárquico 1 está la constitución." .
:normHierarchy rdfs:domain ludlaw:Norma .
:normHierarchy rdfs:range xsd:string .

Lenguaje
:lenguaje rdf:type rdf:Property .
:lenguaje rdfs:label "Lenguaje" .
:lenguaje rdfs:comment "Idioma en el que la norma está escrita" .
:lenguaje owl:equivalentProperty dc:language .
:lenguaje rdfs:domain ludlaw:Norma .
:lenguaje rdfs:range xsd:string .

Name President
:namePresident rdf:type rdf:Property .
:namePresident rdfs:label "Name President" .
:namePresident rdfs:comment "Nombre del presidente cuando fue creada la norma." .
:namePresident owl:equivalentProperty foaf:name .
:namePresident rdfs:domain ludlaw:Norma .
:namePresident rdfs:range xsd:string .

Creator
:creator rdf:type rdf:Property .
:creator rdfs:label "Creator" .
:creator rdfs:comment "El creador de las normas serán las funciones del estado como: la Función
Ejecutiva." .
:creator owl:equivalentProperty dc:creator .
:creator rdfs:domain ludlaw:Norma .
:creator rdfs:range xsd:string .

Name Register
:nameRegister rdf:type rdf:Property .
:nameRegister rdfs:label "Name Register" .
:nameRegister rdfs:comment "Es el nombre que se asigna cada país al diario oficial donde se
publican las normas." .
:nameRegister owl:equivalentProperty foaf:name .
:nameRegister rdfs:domain ludlaw:Norma .
:nameRegister rdfs:range xsd:string .

Number Register
:numberRegister rdf:type rdf:Property .
:numberRegister rdfs:label "Number Register" .
:numberRegister rdfs:comment "Es el número del registro oficial en el que ha sido publicada la
norma." .

```

```

:numberRegister rdfs:domain ludlaw:Norma .
:numberRegister rdfs:range xsd:string .
Url Norm
:urlNorm rdf:type rdf:Property .
:urlNorm rdfs:label "Url Norm" .
:urlNorm rdfs:comment "El identificador uniforme de recursos utilizado para identificar una
norma." .
:urlNorm owl:equivalentProperty foaf:weblog .
:urlNorm rdfs:domain ludlaw:Norma .
:urlNorm rdfs:range rdfs:resource .
Composite
:composite rdf:type rdf:Property .
:composite rdfs:label "Composite" .
:composite rdfs:comment "Define la relacion entre norma y articulo." .
:composite rdfs:domain ludlaw:Norma .
:composite rdfs:range ludlaw:Articulo .
Is Part Of
:isPartOf rdf:type rdf:Property .
:isPartOf rdfs:label "Is part of" .
:isPartOf rdfs:comment "Define la relacion entre articulo y norma." .
:isPartOf rdfs:domain ludlaw:Articulo .
:isPartOf rdfs:range ludlaw:Norma .
Organism
:Organism rdf:type rdfs:class .
:Organism rdfs:label "Organism" .
:Organism rdfs:comment "La clase Organismo contiene información de los organismos que rigen el
sistema de educación superior." .
:Organism rdfs:subclassOf foaf:Organization .
:Organism owl:equivalentClass aiiso:Institution, dbpedia-owl:Organisation, dbpedia-
owl:University, schema:CollegeOrUniversity, schema:Organization, schema:EducationalOrganization
.
Organism Name
:orgName rdf:type rdf:Property .
:orgName rdfs:label "Organism Name" .

:orgName rdfs:comment "Define el nombre de los organismos que rigen el Sistema de Educación
Superior." .
:orgName owl:equivalentProperty foaf:name .
:orgName rdfs:domain ludlaw:Organismo .
:orgName rdfs:range xsd:string .
Organism Description
:orgDescription rdf:type rdf:Property .
:orgDescription rdfs:label "Organism Description " .
:orgDescription rdfs:comment "Es una breve descripción de los organismos de educación Superior."
.

:orgDescription owl:equivalentProperty dc:description .
:orgDescription rdfs:domain ludlaw:Organismo .
:orgDescription rdfs:range xsd:string .
Organism Acronym
:orgAcronym rdf:type rdf:Property .
:orgAcronym rdfs:label "Organism Acronym" .

```

```

:orgAcronym rdfs:comment "Son las iniciales del nombre de los organismos." .
:orgAcronym rdfs:domain ludlaw:Organismo .
:orgAcronym rdfs:range xsd:string .
Founding Date
:foundingDate rdf:type rdf:Property .
:foundingDate rdfs:label "Founding Date" .
:foundingDate owl:equivalentProperty schema:foundingDate .
:foundingDate rdfs:subpropertyOf dc:date .
:foundingDate rdfs:domain ludlaw:Organismo .
:foundingDate rdfs:range xsd:string .
Homepage
:homepage rdf:type rdf:Property .
:homepage rdfs:label "Homepage" .
:homepage rdfs:comment "Es la página web del organismo de Educación Superior." .
:homepage owl:equivalentProperty foaf:homepage .
:homepage rdfs:subpropertyOf foaf:page .
:homepage rdfs:domain ludlaw:Organismo .
:homepage rdfs:range xsd:string .
Organism Type
:orgType rdf:type rdf:Property .

:orgType rdfs:label "Organism Type" .
:orgType rdfs:comment "Se clasifican los organismos de educación superior encategorías como:
Organismo regulador, Organismo de acreditacion." .
:orgType rdfs:domain ludlaw:Organismo .
:orgType rdfs:range xsd:string .
Organismo Logo
:orgLogo rdf:type rdf:Property .
:orgLogo rdfs:label "Organism Logo" .
:orgLogo rdfs:comment "Contiene un enlace que muestra el logotipo del organismo." .
:orgLogo rdfs:domain ludlaw:Organismo .
:orgLogo rdfs:range rdfs:resource .
Article
:Article rdf:type rdfs:class .
:Article rdfs:label "Article" .
:Article rdfs:comment "Contiene cada uno de los artículos como: número, descripción, categoría y
tags." .
Number Article
:numberArticle rdf:type rdf:Property .
:numberArticle rdfs:label "Number Article" .
:numberArticle rdfs:comment "Es el número del artículo de cada uno de las normas." .
:numberArticle rdfs:domain ludlaw:Articulo .
:numberArticle rdfs:range xsd:string .
Text Article
:textArticle rdf:type rdf:Property .
:textArticle rdfs:label "Text Article" .
:textArticle rdfs:comment "Contiene informacion que es parte de un tratado o documento oficial
de una norma." .
:textArticle rdfs:domain ludlaw:Articulo .
:textArticle rdfs:range xsd:string .
Description Article

```

```

:descriptionArticle rdf:type rdf:Property .
:descriptionArticle rdfs:label "Description Article" .
:descriptionArticle rdfs:comment "Una breve descripción de cada uno de los :artículos." .
:descriptionArticle owl:equivalentProperty dc:description .
:descriptionArticle rdfs:domain ludlaw:Articulo .
:descriptionArticle rdfs:range xsd:string .

Tag
:tag rdf:type rdf:Property .
:tag rdfs:label "Tag" .
:tag rdfs:comment "Son palabras claves que describen el contenido del artículo." .
:tag rdfs:domain ludlaw:Articulo .
:tag rdfs:range xsd:string .

University
:University rdf:type rdfs:class .
:University rdfs:label "University" .
:University rdfs:comment "La universidad es una institución educativa de tercer nivel." .
:University owl:equivalentClass aiiso:Institution, dbpedia-owl:Organisation, dbpedia-owl:University, schema:CollegeOrUniversity, schema:Organization, schema:EducationalOrganization
.

University Name
:uniName rdf:type rdf:Property .
:uniName rdfs:label "University Name" .
:uniName rdfs:comment "Se define el nombre de la universidad de cada país." .
:uniName rdfs:domain ludlaw:Universidad .
:uniName rdfs:range xsd:string .

University Uri
:uniUri rdf:type rdf:Property .
:uniUri rdfs:label "University Uri" .
:uniUri rdfs:comment "La uri de DBPEDIA de las universidades." .
:uniUri rdfs:domain ludlaw:Universidad .
:uniUri rdfs:range rdfs:resource .

```


Anexo 5. Formato RDF generado mediante Jena

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:foaf="http://xmlns.com/foaf/0.1/"
  xmlns:ludlaw="http://utpl.edu.ec/educacion/leyes/ludlaw/"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#" >
<rdf:Description rdf:about="http://utpl.edu.ec/educacion/leyes/Article/ART18">
<ludlaw:isPartOf rdf:resource="http://utpl.edu.ec/educacion/leyes/Norm/NORMEC01"/>
<rdf:type rdf:resource="http://www.w3.org/2000/01/rdf-schema#Class"/>
<dc:subject>asignaciones presupuestarias, ley</dc:subject>
<dc:description>FINANCIAMIENTO UNIVERSITARIO</dc:description>
<ludlaw:textArticle>Privación de rentas.- La Función Ejecutiva no
podrá privar de sus rentas o asignaciones presupuestarias, o
retardar las transferencias a ninguna institución del sistema,
salvo en los casos previstos en esta Ley.</ludlaw:textArticle>
<rdfs:range
rdf:datatype="java:com.hp.hpl.jena.rdf.model.impl.ResourceImpl">http://www.w3.org/200
1/XMLSchema#string</rdfs:range>
<ludlaw:numberArticle>Art. 22</ludlaw:numberArticle>
</rdf:Description>
<rdf:Description rdf:about="http://utpl.edu.ec/educacion/leyes/Article/ART68">
<ludlaw:isPartOf rdf:resource="http://utpl.edu.ec/educacion/leyes/Norm/NORMEC01"/>
<rdf:type rdf:resource="http://www.w3.org/2000/01/rdf-schema#Class"/>
<dc:subject>profesor, título profesional, sistema de educación superior</dc:subject>
<dc:description>PERSONAL ACADÉMICO</dc:description>
<ludlaw:textArticle>Profesor o profesora titular en institutos superiores y
conservatorios
superiores.- Para ser profesor
o profesora titular de un instituto superior técnico, tecnológico, de artes o
conservatorio superior se requiere
tener un título profesional y demás requisitos que establezca el Reglamento de
Carrera y Escalafón del Profesor e
Investigador del Sistema de Educación Superior.</ludlaw:textArticle>
<rdfs:range
rdf:datatype="java:com.hp.hpl.jena.rdf.model.impl.ResourceImpl">http://www.w3.org/200
1/XMLSchema#string</rdfs:range>
<ludlaw:numberArticle>Art. 154</ludlaw:numberArticle>
</rdf:Description>
<rdf:Description
rdf:about="http://utpl.edu.ec/educacion/leyes/Article/ART120"><ludlaw:isPartOf
rdf:resource="http://utpl.edu.ec/educacion/leyes/Norm/NORMME01"/>
<rdf:type rdf:resource="http://www.w3.org/2000/01/rdf-schema#Class"/>
<dc:subject>reconocimiento de validez oficial, incumplimiento, clausurar el servicio
educativo</dc:subject>
```

```

<dc:description>TITULOS ACADEMICOS</dc:description>
<ludlaw:textArticle>Los particulares que impartan estudios de tipo superior con
autorización o reconocimiento de validez oficial deberán registrarse en la Secretaría
de Educación Pública. El incumplimiento de esta disposición motivará la imposición de
multa hasta de cien mil pesos, y en caso de persistir el incumplimiento se podrá
clausurar el servicio educativo
</ludlaw:textArticle>
<rdfs:range
rdf:datatype="java:com.hp.hpl.jena.rdf.model.impl.ResourceImpl">http://www.w3.org/200
1/XMLSchema#string</rdfs:range>
<ludlaw:numberArticle>Art. 19</ludlaw:numberArticle>
</rdf:Description>
<rdf:Description rdf:about="http://utpl.edu.ec/educacion/leyes/Article/ART88">
<ludlaw:isPartOf
rdf:resource="http://utpl.edu.ec/educacion/leyes/Norm/NORMAR01"/>
<rdf:type rdf:resource="http://www.w3.org/2000/01/rdf-schema#Class"/>
<dc:subject>instituciones universitarias, reformar estatutos, investigación
científica, autonomía académica, ética profesional</dc:subject>
<dc:description>AUTONOMIA UNIVERSITARIA</dc:description>
<ludlaw:textArticle>Las instituciones universitarias tendrán autonomía académica e
institucional, que comprende básicamente las siguientes atribuciones:
a) Dictar y reformas sus estatutos, los que serán comunicados al Ministerio de
Educación a los fines establecidos en el artículo 34 de la presente ley;
b) Definir sus órganos de gobierno, establecer sus funciones, decidir su integración
y elegir sus autoridades de acuerdo a lo que establezcan los estatutos y lo que
prescribe la presente ley;c) Administrar sus bienes y recursos, conforme a sus
estatutos y las leyes que regulan la materia;
d) Crear carreras universitarias de grado y de posgrado;
e) Formular y desarrollar planes de estudio, de investigación científica y de
extensión y servicios a la comunidad incluyendo la enseñanza de la ética profesional;
f) Otorgar grados académicos y títulos habilitantes conforme a las condiciones que se
establezcan en la presente ley;
g) Impartir enseñanza, con lo fines de experimentación, de innovación pedagógica o de
práctica profesional docente, en los niveles preuniversitarios, debiendo continuar en
funcionamiento los establecimientos existentes actualmente que reúnan dichas
características;
h) Establecer el régimen de acceso, permanencia y promoción del personal docente y no
docente;

i) Designar y remover al personal;j) Establecer el régimen de admisión, permanencia y
promoción de los estudiantes, así como el régimen de equivalencias;
k) Revalidar, sólo como atribución de las universidades nacionales, títulos
extranjeros;
l) Fijar el régimen de convivencia;
m) Desarrollar y participar en emprendimientos que favorezcan el avance y aplicación
de los conocimientos;

```

n) Mantener relaciones de carácter educativo, científico y cultural con instituciones del país y del extranjero;

ñ) Reconocer oficialmente asociaciones de estudiantes, cumplidos que sean los requisitos que establezca la reglamentación, lo que conferirá a tales entidades personería jurídica.

```

</ludlaw:textArticle>
<rdfs:range
rdf:datatype="java:com.hp.hp1.jena.rdf.model.impl.ResourceImpl">http://www.w3.org/200
1/XMLSchema#string</rdfs:range>
<ludlaw:numberArticle>Art. 29</ludlaw:numberArticle>
</rdf:Description>
<rdf:Description rdf:about="http://utpl.edu.ec/educacion/leyes/Norm/NORMAR02">
<ludlaw:numberRegister>S/I</ludlaw:numberRegister>
<ludlaw:promulgationDate>1994/08/22</ludlaw:promulgationDate>
<rdf:type rdf:resource="http://www.w3.org/2000/01/rdf-schema#Class"/>
<ludlaw:normType>Constitución</ludlaw:normType>
<ludlaw:publicationDate>1994/08/22</ludlaw:publicationDate>
<dc:title>Constitución de La nación Argentina</dc:title>
<ludlaw:nameRegister>Boletín Oficial </ludlaw:nameRegister>
<rdfs:range
rdf:datatype="java:com.hp.hp1.jena.rdf.model.impl.ResourceImpl">http://www.w3.org/200
1/XMLSchema#date</rdfs:range>
<foaf:name>Carlos Saúl Menem </foaf:name>
<rdfs:range
rdf:datatype="java:com.hp.hp1.jena.rdf.model.impl.ResourceImpl">http://www.w3.org/200
1/XMLSchema#string</rdfs:range>
<dc:language>Español</dc:language>
<ludlaw:normHierarchy>Nivel 1: Constitución</ludlaw:normHierarchy>
<dc:creator>Función Legislativa</dc:creator>
<ludlaw:urlNorm>http://www.constitution.org/cons/argentin.htm</ludlaw:urlNorm>
</rdf:Description>
<rdf:Description rdf:about="http://utpl.edu.ec/educacion/leyes/Article/ART51">
<ludlaw:isPartOf rdf:resource="http://utpl.edu.ec/educacion/leyes/Norm/NORMEC01"/>
<rdf:type rdf:resource="http://www.w3.org/2000/01/rdf-schema#Class"/>
<dc:subject>igualdad de oportunidad, paridad de género, autoridades académicas,
equidad</dc:subject>

<dc:description>AUTORIDADES ACADÉMICAS</dc:description>
<ludlaw:textArticle>Paridad de género, igualdad de oportunidades y
equidad.- Cuando existan listas para la elección de rector o
rectora, vicerrector o vicerrectora, vicerrectores o
vicerrectoras, y demás autoridades académicas, deberán ser
integradas respetando la alternancia, la paridad de género,
igualdad de oportunidades y equidad conforme a la
Constitución.</ludlaw:textArticle><rdfs:range

```

```

rdf:datatype="java:com.hp.hpl.jena.rdf.model.impl.ResourceImpl">http://www.w3.org/200
1/XMLSchema#string</rdfs:range>
<ludlaw:numberArticle>Art. 56</ludlaw:numberArticle>
</rdf:Description>
<rdf:Description rdf:about="http://utpl.edu.ec/educacion/leyes/Article/ART114">
<ludlaw:isPartOf rdf:resource="http://utpl.edu.ec/educacion/leyes/Norm/NORMME01"/>
<rdf:type rdf:resource="http://www.w3.org/2000/01/rdf-schema#Class"/>
<dc:subject>sistema de educación superior, planeación institucional, superación
académica</dc:subject>
<dc:description>FINANCIAMIENTO UNIVERSITARIO</dc:description><ludlaw:textArticle>Los
recursos que conforme al Presupuesto de Egresos de la Federación se asignen a las
instituciones de educación superior se
determinarán atendiendo a las prioridades nacionales y a la participación de las
instituciones en el desarrollo del sistema de educación superior y considerando la
planeación institucional y los programas de superación académica y de mejoramiento
administrativo, así como el conjunto de gastos de operación previstos. Para decidir
la asignación de los recursos a que se refiere el párrafo

anterior, en ningún caso se tomarán en cuenta consideraciones ajenas a las
educativas. </ludlaw:textArticle>
<rdfs:range
rdf:datatype="java:com.hp.hpl.jena.rdf.model.impl.ResourceImpl">http://www.w3.org/200
1/XMLSchema#string</rdfs:range>
<ludlaw:numberArticle>Art. 23</ludlaw:numberArticle>
</rdf:Description>
<rdf:Description rdf:about="http://utpl.edu.ec/educacion/leyes/Article/ART39">
<ludlaw:isPartOf rdf:resource="http://utpl.edu.ec/educacion/leyes/Norm/NORMME01"/>
<rdf:type rdf:resource="http://www.w3.org/2000/01/rdf-schema#Class"/>
<dc:subject>publicación de información, sistema de educación superior,
remuneraciones</dc:subject>
<dc:description>FINANCIAMIENTO UNIVERSITARIO</dc:description>
<ludlaw:textArticle>Publicación de información en portalelectrónico.- Las
instituciones que conforman el Sistema de
Educación Superior, en cumplimiento de la Ley, obligatoriamente deberán publicar en su
portal electrónico las remuneraciones de sus autoridades, profesores,
investigadores, servidores y trabajadores.
Esta información se integrará de manera obligatoria al Sistema Nacional de Información
de la Educación Superior del Ecuador.</ludlaw:textArticle>
<rdfs:range
rdf:datatype="java:com.hp.hpl.jena.rdf.model.impl.ResourceImpl">http://www.w3.org/200
1/XMLSchema#string</rdfs:range>
<ludlaw:numberArticle>Art. 43</ludlaw:numberArticle>
</rdf:Description>
</rdf:RDF>

```

Anexo6. Instalación de Virtuoso

Para poder descargar Virtuoso se puede ingresar a la siguiente dirección: <http://download.openlinksw.com/> , cuando visite el sitio web, seleccione "Descargar y evaluación".

Luego se debe seguir algunos pasos:

- ✓ Seleccione el producto "Virtuoso Universal Server"
- ✓ De "sistema operativo de servidor", seleccione "Windows" y haga clic en "Siguiente"
- ✓ Seleccione una versión y haga clic en "Siguiente"
- ✓ Luego empieza la descarga, una vez terminada ya tenemos el instalador listo para empezar la instalación.
- ✓ Nos aparece la pantalla de bienvenida

En esta pantalla podemos observar que la instalación está en proceso:

Después de la instalación del servidor de Virtuoso puede iniciarse mediante la ejecución del "Virtuoso Service Manager", que se añade en la barra de estado. Hacer doble clic en el "Virtuoso Service Manager", para cargar la aplicación, seleccione la opción "OpenLink Virtuoso Server" y pulse el botón "Inicio" para ejecutar el servicio.

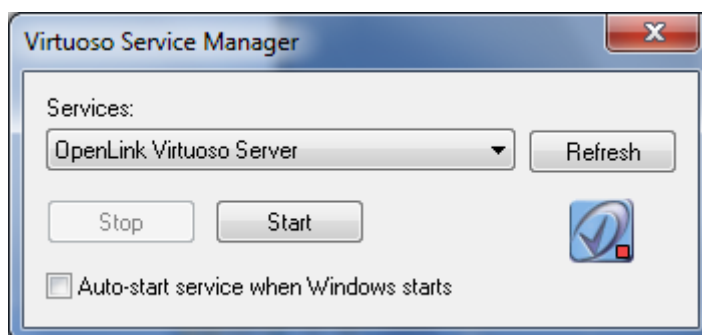


Ilustración 40 Inicializar el servicio Virtuoso

Una vez que el servicio de Virtuoso ha iniciado con éxito. Se coloca en el navegador <http://localhost:8890> y nos presenta la página de bienvenida indicando que el servidor está en línea.

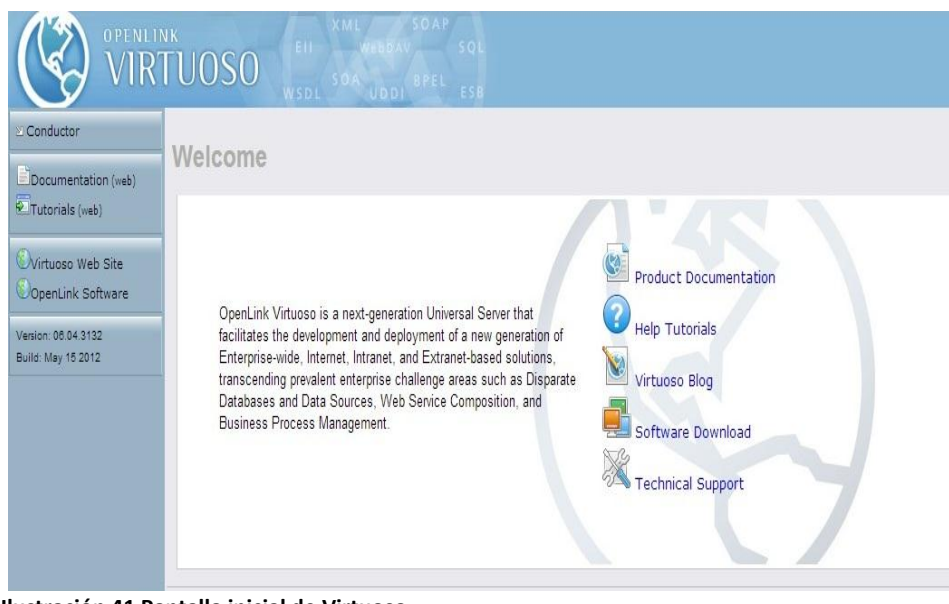


Ilustración 41 Pantalla inicial de Virtuoso

Para logearse hacemos clic en conductor donde nos va a pedir el account y password que por defecto es dba dba.



Ilustración 42 Login Virtuoso

Anexo 7. Cargar archivo RDF en Virtuoso

Para realizar la carga del archivo RDF se hace clic en la pestaña Linked Data y luego seleccionamos Quad Store Upload como se puede observar en la ilustración 43.

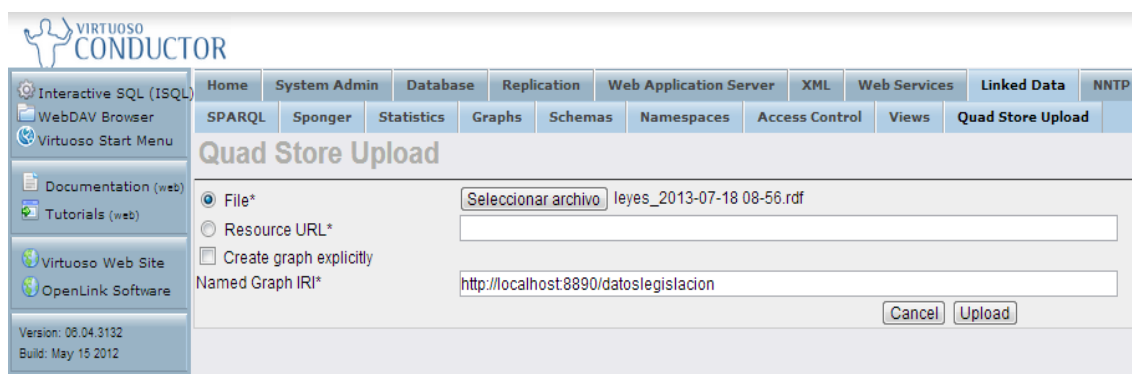


Ilustración 43 Carga del RDF

Luego se escoge el archivo RDF y se le asigna un nombre al grafo en este caso se llama <http://utp.edu.ec/educacion/superior/datoslegislacion>, se da clic en Upload y listo el grafo puede ser consultado. En la ilustración 44 podemos observar los grafos que han sido almacenados para luego ser procesados mediante consultas SPARQL.



Ilustración 44 Grafos RDF

Virtuoso nos permite agregar los prefijos con los cuales se va a trabajar algunos ya vienen previamente añadidos, en la ilustración 45 se puede observar como añadir un nuevo prefijo.

Prefix	URI
bif	bif.
dawgt	http://www.w3.org/2001/sw/DataAccess/tests/test-dawg#
dbpedia	http://dbpedia.org/resource/
dbpprop	http://dbpedia.org/property/
dc	http://purl.org/dc/elements/1.1/
fn	http://www.w3.org/2005/xpath-functions/#
foaf	http://xmlns.com/foaf/0.1/
geo	http://www.w3.org/2003/01/geo/wgs84_pos#
go	http://purl.org/obo/owl/GO#
math	http://www.w3.org/2000/10/swap/math#
mesh	http://purl.org/commons/record/mesh/
mf	http://www.w3.org/2001/sw/DataAccess/tests/test-manifest#
nci	http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#
obo	http://www.geneontology.org/formats/oboInOwl#
owl	http://www.w3.org/2002/07/owl#
product	http://www.buy.com/rss/module/productV2/
protseq	http://purl.org/science/protein/bysequence/
rdf	http://www.w3.org/1999/02/22-rdf-syntax-ns#
rdfa	http://www.w3.org/ns/rdfa#
rdfdf	http://www.openlinksw.com/virtrdf-data-formats#

Prefix:

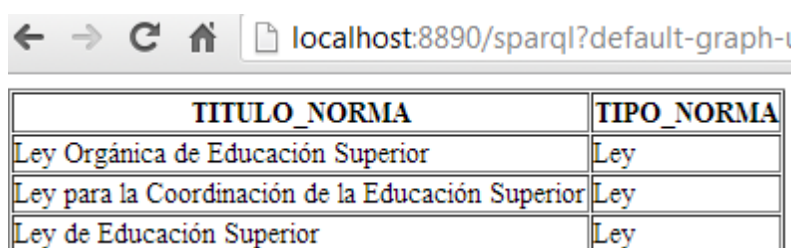
URI:

Ilustración 45 Añadir Prefijos en Virtuoso

Anexo 8. Consultas SPARQL

¿Cuáles son las normas de educación superior que son de tipo de ley?

```
SELECT ?TITULO_NORMA ?TIPO_NORMA
WHERE
{
  ?NORMAS dc:title ?TITULO_NORMA.
  ?NORMAS ludlaw:normType ?TIPO_NORMA.
  FILTER regex(?TIPO_NORMA, "Ley")
}
```

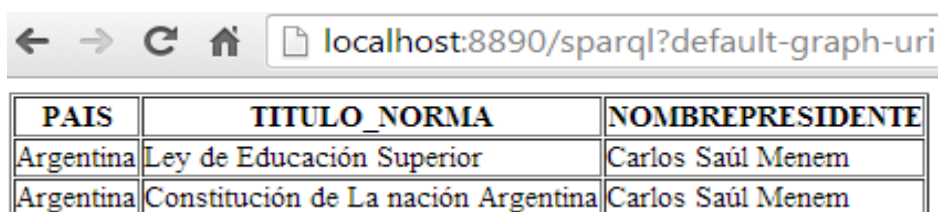


TITULO_NORMA	TIPO_NORMA
Ley Orgánica de Educación Superior	Ley
Ley para la Coordinación de la Educación Superior	Ley
Ley de Educación Superior	Ley

Ilustración 46 Resultado Tipo de Norma

¿Cuáles es el nombre del presidente que en su administración se crearon las normas de educación superior de Argentina?

```
SELECT ?PAIS ?TITULO_NORMA ?NOMBREPRESIDENTE
WHERE
{
  ?SISTEMAEDU ludlaw:regulatedBy ?NORMA.
  ?PAISES ludlaw:hasSystem ?SISTEMAEDU.
  ?SISTEMAEDU ludlaw:systemEducation ?SISTEMA_EDUCACION.
  ?PAISES ludlaw:nameCountry ?PAIS.
  ?NORMA dc:title ?TITULO_NORMA.
  ?NORMA foaf:name ?NOMBREPRESIDENTE.
  FILTER regex(?PAIS, "Argentina")
}
```

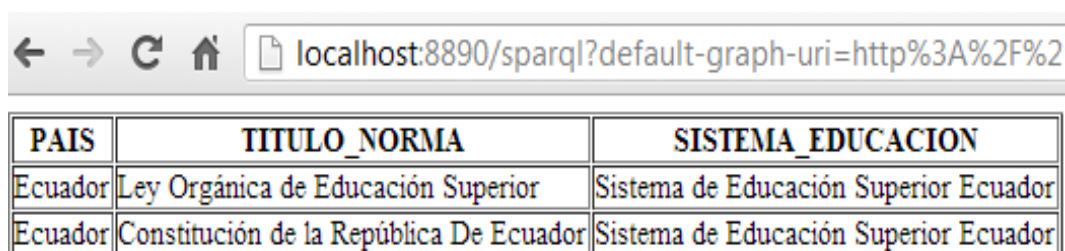


PAIS	TITULO_NORMA	NOMBREPRESIDENTE
Argentina	Ley de Educación Superior	Carlos Saúl Menem
Argentina	Constitución de La nación Argentina	Carlos Saúl Menem

Ilustración 47 Resultado Normas y Nombre del Presidente

¿Qué normas X regulan el sistema de educación Superior en Ecuador?

```
SELECT ?PAIS ?TITULO_NORMA ?SISTEMA_EDUCACION
WHERE
{
  ?SISTEMAEDU ludlaw:regulatedBy ?NORMA.
  ?PAISES ludlaw:hasSystem ?SISTEMAEDU.
  ?SISTEMAEDU ludlaw:systemEducation ?SISTEMA_EDUCACION.
  ?PAISES ludlaw:nameCountry ?PAIS.
  ?NORMA dc:title ?TITULO_NORMA
  FILTER REGEX(?PAIS, "Ecuador")
}
```



PAIS	TITULO_NORMA	SISTEMA_EDUCACION
Ecuador	Ley Orgánica de Educación Superior	Sistema de Educación Superior Ecuador
Ecuador	Constitución de la República De Ecuador	Sistema de Educación Superior Ecuador

Ilustración 48 Resultado Normas que Regulan Sistema Educación

¿Cuáles son las normas que en su articulado mencionan sobre el financiamiento universitario?

```
SELECT ?TITULO_NORMA ?NUMERO_ARTICULO ?TEXTO_ARTICULO ?CATEGORIA
WHERE
{
  ?ARTICULO ludlaw:isPartOf ?NORMA.
  ?NORMA dc:title ?TITULO_NORMA
  FILTER REGEX(?ARTICULO, "http://utpl.edu.ec/educacion/leyes/ludlaw/Article/")
  FILTER REGEX(?NORMA, "http://utpl.edu.ec/educacion/leyes/ludlaw/Norm/")
  ?ARTICULO ludlaw:numberArticle ?NUMERO_ARTICULO.
  ?ARTICULO ludlaw:textArticle ?TEXTO_ARTICULO.
  ?ARTICULO dc:description ?CATEGORIA
  FILTER REGEX(?CATEGORIA, "FINANCIAMIENTO UNIVERSITARIO")
  ?SISTEMAEDU ludlaw:regulatedBy ?NORMA.
  FILTER REGEX(?SISTEMAEDU,
"http://utpl.edu.ec/educacion/leyes/ludlaw/SistemaEducacion/")
  ?PAISES ludlaw:hasSystem ?SISTEMAEDU.
  ?PAISES ludlaw:nameCountry ?PAIS
}
```

TITULO_NORMA	NUMERO_ARTICULO	TEXTO_ARTICULO	CATEGORIA
Ley de Educación Superior	Art. 60	Las instituciones universitarias nacionales podrán promover la constitución de fundaciones, sociedades u otras formas de asociación civil, destinada a apoyar su labor, a facilitar las relaciones con el medio, a dar respuesta a sus necesidades y a promover las condiciones necesarias para el cumplimiento de sus fines y objetivos.	FINANCIAMIENTO UNIVERSITARIO
Ley de Educación Superior	Art. 59	Las instituciones universitarias nacionales tienen autarquía económico-financiera, la que ejercerán dentro del régimen de la ley 24.156 de Administración Financiera y Sistemas de Control del Sector Público Nacional. En ese marco corresponde a dichas instituciones: a) Administrar su patrimonio y aprobar su presupuesto. Los recursos no utilizados al cierre de cada ejercicio, se transferirán automáticamente al siguiente; b) Fijar su régimen salarial y de administración de personal; c) Podrán dictar normas relativas a la generación de recursos adicionales a los aportes del Tesoro nacional, mediante la venta de bienes, productos, derechos o servicios, subsidios, contribuciones, herencias, derechos o tasas por los servicios que presten, así como todo otro recurso que pudiera corresponderles por cualquier título o actividad. Los recursos adicionales que provienen de contribuciones o tasas por los estudios de grado, deberán destinarse prioritariamente a becas, préstamos, subsidios o créditos u otro tipo de ayuda estudiantil y apoyo docente; estos recursos adicionales no podrán utilizarse para financiar gastos corrientes. Los sistemas de becas, préstamos u otro tipo de ayuda estarán fundamentalmente destinados a aquellos estudiantes que demuestren aptitud suficiente y respondan adecuadamente a las exigencias académicas de la institución y que por razones económicas no pudieran acceder o continuar los estudios universitarios, de forma tal que nadie se vea imposibilitado por ese motivo de cursar tales estudios; d) Garantizar el normal desenvolvimiento de sus unidades asistenciales, asegurándose el manejo descentralizado de los fondos que ellas generen, con acuerdo a las normas que dicten sus Consejos Superiores y a la legislación vigente; e) Constituir personas jurídicas de derecho público o privado, o participar en ellas, no requiriéndose adoptar una forma jurídica diferente para acceder a los beneficios de la ley 23.877; f) Aplicar el régimen general de contrataciones, de responsabilidad patrimonial y de gestión de bienes reales, con las excepciones que establezca la legislación. El rector y los miembros del Consejo Superior de las instituciones universitarias nacionales serán responsables de su administración según su participación, debiendo responder en los términos y con los alcances previstos en los artículos 130 y 131 de la ley 24.156. En ningún caso el Estado nacional responderá por las obligaciones asumidas por las instituciones universitarias que importen un perjuicio para el Tesoro nacional.	FINANCIAMIENTO UNIVERSITARIO
Ley de Educación Superior	Art. 61	El Congreso Nacional debe disponer de la partida presupuestaria anual correspondiente al nivel de educación superior, de un porcentaje que será destinado a becas y subsidios en ese nivel.	FINANCIAMIENTO UNIVERSITARIO
Ley de Educación Superior	Art. 58	Corresponde al Estado nacional asegurar el aporte financiero para el sostenimiento de las instituciones universitarias nacionales, que garantice su normal funcionamiento, desarrollo y cumplimiento de sus fines. Para la distribución de ese aporte entre las mismas se tendrán especialmente en cuenta indicadores de eficiencia y equidad. En ningún caso podrá disminuirse el aporte del Tesoro nacional como contrapartida de la generación de recursos complementarios por parte de las instituciones universitarias nacionales.	FINANCIAMIENTO UNIVERSITARIO
Ley Orgánica de Educación Superior	Art. 22	Privación de rentas. - La Función Ejecutiva no podrá privar de sus rentas o asignaciones presupuestarias, o retardar las transferencias a ninguna institución del sistema, salvo en los casos previstos en esta Ley.	FINANCIAMIENTO UNIVERSITARIO
Ley Orgánica de Educación Superior	Art. 43	Publicación de información en portal electrónico. - Las instituciones que conforman el Sistema de Educación Superior, en cumplimiento de la Ley, obligatoriamente deberán publicar en su portal electrónico las remuneraciones de sus autoridades, profesores, investigadores, servidores y trabajadores. Esta información se integrará de manera obligatoria al Sistema Nacional de Información de la Educación Superior del Ecuador.	FINANCIAMIENTO UNIVERSITARIO
Ley para la Coordinación de la Educación Superior	Art. 23	Los recursos que conforme al Presupuesto de Egresos de la Federación se asignen a las instituciones de educación superior se determinarán atendiendo a las prioridades nacionales y a la participación de las instituciones en el desarrollo del sistema de educación superior y considerando la planeación institucional y los programas de superación académica y de mejoramiento administrativo, así como el conjunto de gastos de operación previstos. Para decidir la asignación de los recursos a que se refiere el párrafo anterior, en ningún caso se tomarán en cuenta	FINANCIAMIENTO UNIVERSITARIO

Ilustración 49 Resultado Normas sobre Financiamiento Universitario

¿Cuáles son los organismos que regulan la educación superior en el país de Ecuador?

```
SELECT ?PAIS ?NOMBRE_ORGANISMO ?TIPO_ORGANISMO ?SIGLAS ?HOMEPAGE
WHERE {
?ORGANISMOS ludlaw:established ?NORMA.
?NORMA dc:title ?TITULO_NORMA.
?ORGANISMOS foaf:name ?NOMBRE_ORGANISMO.
?ORGANISMOS ludlaw:orgType ?TIPO_ORGANISMO.
?ORGANISMOS ludlaw:orgAcronym ?SIGLAS.
?ORGANISMOS foaf:homepage ?HOMEPAGE.
?SISTEMAEDU ludlaw:regulatedBy ?NORMA.
FILTER REGEX(?SISTEMAEDU,
"http://utpl.edu.ec/educacion/leyes/ludlaw/SistemaEducacion/")
?PAISES ludlaw:hasSystem ?SISTEMAEDU.
?PAISES ludlaw:nameCountry ?PAIS
FILTER REGEX(?PAIS,"Ecuador")
}
```

localhost:8890/sparql?default-graph-uri=http%3A%2Futpl.edu.ec%2Feducacion%2Fleyes%2Fludlaw%2F&query=SELECT+%3

PAIS	NOMBRE_ORGANISMO	TIPO_ORGANISMO	SIGLAS	HOMEPAGE
Ecuador	Consejo de Evaluación, Acreditación y Aseguramiento de la Calidad de la Educación Superior	Organismo de Evaluación y Acreditación	CEAACES	http://www.ceaaces.gob.ec
Ecuador	Consejo de Educación Superior	Organismo de Coordinación y Consulta	CES	http://www.ces.gob.ec

Ilustración 50 Resultado Organismos reguladores

¿Cuáles son los organismos encargados de la coordinación y consulta, una breve descripción y cual es la fecha de fundación en los diferentes países?

```

SELECT ?PAIS ?NOMBRE_ORGANISMO ?SIGLAS ?DESCRIPCION_ORGANISMO ?FUNDACION
WHERE {
?ORGANISMOS ludlaw:established ?NORMA.
?NORMA dc:title ?TITULO_NORMA.
?ORGANISMOS foaf:name ?NOMBRE_ORGANISMO.
?ORGANISMOS ludlaw:orgType ?TIPO_ORGANISMO.
    FILTER REGEX(?TIPO_ORGANISMO,"Organismo de Coordinación y Consulta")
?ORGANISMOS ludlaw:orgAcronym ?SIGLAS.
?ORGANISMOS dc:description ?DESCRIPCION_ORGANISMO.
?ORGANISMOS ludlaw:foundingDate ?FUNDACION.
?SISTEMAEDU ludlaw:regulatedBy ?NORMA.
    FILTER REGEX(?SISTEMAEDU,
"http://utpl.edu.ec/educacion/leyes/ludlaw/SistemaEducacion/")
?PAISES ludlaw:hasSystem ?SISTEMAEDU.
?PAISES ludlaw:nameCountry ?PAIS
}

```

PAIS	NOMBRE_ORGANISMO	DESCRIPCION_ORGANISMO	FUNDACION
Ecuador	Consejo de Educación Superior	Es el organismo de derecho público con personería jurídica, con patrimonio propio, independencia administrativa, financiera y operativa, que tiene por objetivo la planificación, regulación y coordinación interna del Sistema de Educación Superior, y la relación entre sus distintos actores con la Función Ejecutiva y la sociedad ecuatoriana.	2010
Argentina	Consejo Interuniversitario Nacional	Esencialmente, de coordinación de políticas universitarias y promoción de políticas y actividades de interés para el sistema público de Educación Superior. Es, además, órgano de consulta obligada en la toma de decisiones de trascendencia para el sistema universitario.	1985
México	Consejo Nacional Consultivo de Educación Normal	Es el órgano de consulta de la Secretaría de Educación Pública, de las entidades federativas cuando éstas lo soliciten y de las instituciones de educación normal para coordinar sus actividades, orientar la celebración de los convenios que sobre la materia prevé esta ley y contribuir a vincular dicha educación con los requerimientos del país, de conformidad con la política educativa nacional.	1970

Ilustración 51 Resultado Organismos de Consulta

Publicación de datos enlazados en el ámbito de la Legislación de Educación Superior en Latinoamérica

María Fernanda Valverde¹, Juan Carlos Morocho¹ y Nelson Piedra¹

¹ Tecnologías Avanzadas de la Web y Sistemas Basados en el Conocimiento
Departamento de Ciencias de la Computación y Electrónica
Universidad Técnica Particular de Loja
{mfvalverde, nopiedra}@utpl.edu.ec

Keywords: Linked Data, RDF, metadatos, ontología, RDF Schema, Web Semántica, URI, XML, SPARQL, Tripletas.

1 Introducción

Con el pasar de los años la Web ha evolucionado a pasos agigantados desde la Web 1.0, Web 2.0 y ahora en la actualidad la Web de Datos Enlazados la cual trata de vincular los distintos datos que se encuentran distribuidos en la Web con el propósito de mejorar la búsqueda de información, lo cual brindará mejor accesibilidad y facilidad de uso de los datos. En diferentes países como Chile, Holanda, Estados Unidos, Reino Unido, Australia, Canadá se están desarrollando proyectos sobre legislación a través de Linked Data, pero el ámbito del proyecto de investigación es sobre Legislación de Educación Superior.

El estudio que se propone es publicar datos del ámbito Legislación de Educación Superior bajo los principios de Linked Open Data (LOD) diseñados por Tim Berners-Lee, de tal manera que estos puedan ser accedidos de forma pública y formen parte de la web de datos para alcanzar el objetivo planteado se utilizará diferentes tecnologías de la Web Semántica como: RDF, SPARQL, motores de almacenamiento de tripletas y herramientas para publicación de datos enlazados de esta manera generar valor agregado a la Web.

2 Contexto

En la actualidad se están realizando varios proyectos sobre web semántica en diferentes áreas de conocimiento algunas como: educación, bioinformática, salud y legislación; basado en este impulso se ha iniciado el proyecto de publicación de datos enlazados en el ámbito de Legislación de Educación Superior de Latinoamérica.

El contexto del estudio de aplicación de Linked Data a la Legislación de Educación Superior en Latinoamérica, la información con la que se cuenta son varios documentos sobre las leyes de Educación Superior, los cuales se encuentran en formatos como son: PDF, Word, Texto plano; información de organismos reguladores, Universidades, normas de Educación Superior que pueden ser leyes, constituciones, decretos entre otras tipos de normas por cada país se analizará cada una de las normas y se obtendrá los artículos de mayor interés. Debido a que la información se encuentra en diferentes fuentes y en formato no estructurado dificulta la extracción de la información; Esta es una limitante a la hora de trabajar en la web de datos, ya

que no permite enlazar la información y aprovechar aquella que ya está publicada y tiene relación con el ámbito que pretendemos publicar.

Trabajos Relacionados

Desde la década de 1940 aparecieron los primeros sistemas de información jurídica (Bizer, Heath, & Berners, *Linked Data - The Story So Far*, 2009), en 1950 los sistemas expertos legales y en la actualidad también se está trabajando en temas jurídicos o gubernamentales que sean de acceso libre, en diferentes países menciona (Cifuentes F. , 2011): Chile, Holanda, Estados Unidos, Reino Unido, Australia, Austria, Canadá, Italia, Nueva Zelanda se están desarrollando proyectos de esta índole. A continuación algunos proyectos relacionados.

En Reino Unido existe el proyecto de *Legislation.gov.uk*, el cual permite a los ciudadanos acceder de forma pública a la información jurídica. Este proyecto abarca las cuatro jurisdicciones las cuales son: Inglaterra, Escocia, Wales e Irlanda del Norte (Casellas, Vallbé, & Bruce, 2011). *Legislation.gov.uk* es un sitio web atractivo, permitiendo un acceso sencillo y directo a la legislación, la gente puede ver hechos enteros, o una sección en particular, ya sea en formato HTML o en versión PDF.

En Estados Unidos, el instituto de Información Legal de la Universidad de Derecho de Cornell tiene como objetivo tratar que la ley sea gratuita, accesible y comprensible.

La Dra. Nuria Casellas participa en proyectos relacionados con Principios de Linked Open Data aplicados al código de regulaciones federales (CFR) el cual se ha publicado en *semanticweb.com*. Según (Casellas N. , *Legally Linked: Linked Open Data Principles Applied To Code Of Federal Regulations*, 2012) en este proyecto los usuarios pueden buscar nombres comerciales de drogas como Tylenol, recibiendo el nombre genérico como un término sugerido. LII Planea incorporar definiciones y vocabularios y la información del producto con la finalidad de que la búsqueda y recuperación de información mejore considerablemente.

También surge la iniciativa en Chile de Datos que promueve prácticas y estándares de publicación de datos públicos. La idea de la organización es que las personas puedan acceder tanto a los datos como a aplicaciones para procesarlos. La Biblioteca del Congreso Nacional tiene avances en Linked Open Data, logrando el nivel 5 en la clasificación de 1 a cinco estrellas que se utiliza para evaluar el nivel de avance en la implementación del sistema, siendo *www.bcn.cl* (Biblioteca Consejo Nacional), la única institución estatal latinoamericana en conseguir este reconocimiento.

3 Proceso para publicación de datos enlazados

Las tecnologías actuales permiten tener una nueva visión de la información, en la que varios elementos (personas, archivos, web's, etc.) están interconectados.

Para dar solución al problema planteado, emplearemos Linked Data, que permite construir la Web de los datos, en una gran base de datos interconectados y distribuidos en la Web. Los datos se enlazan y se examinan de una forma similar a la utilizada para vincular los documentos de tipo HTML.

3.1 Principios de Linked data

Según Berners Lee definió cuatro reglas para la publicación de Linked data (Corcho & Gómez, 2010):

- Usar URIs (Uniform Resource Identifiers) identificando los recursos de forma unívoca;

- Usar URIs http para que la gente pueda acceder a la información del recurso;
- Ofrecer información sobre los recursos usando RDF.
- Incluir enlaces a otros URIs, facilitando el vínculo entre distintos datos distribuidos en la web.

Con la utilización de estos principios lograremos contar con datos interconectados, permite reutilizar la información y de esta manera ofrecer valor agregado a la Web (Corcho & Gómez, 2010).



3.2 Tecnologías

Para la publicación de datos existen diferentes tecnologías las cuales nos permiten realizar el trabajo, algunas de ellas son: RDF, URIS, SPARQL.

- RDF.- es el marco de descripción de recursos para metadatos en la Web elaborado por el W3C. Se basa en la idea de declarar recursos usando la expresión en la forma sujeto-predicado-objeto(Méndez, 2012).
- URIS.- es un identificador único del recurso, es definido por el estándar RFC 3986(Berners-Lee T. , 2005).
- SPARQL.- es un lenguaje de consultas para RDF y en la actualidad es recomendado por W3C.

3.3 Herramientas de almacenamiento de tripletas

En la actualidad existen diferentes herramientas para almacenar tripletas o RDF los cuales se los denomina Triple Store System. Entre las más utilizadas y robustas [9] encontramos las siguientes herramientas: AllegroGraph⁴¹, 4Store⁴², OpenLinkVirtuoso⁴³, Sesame⁴⁴. Para el almacenamiento de las tripletas del proyecto se trabaja con Virtuoso ya que es una herramienta

⁴¹<http://www.franz.com/agraph/allegrograph/>

⁴²<http://4store.org/>

⁴³<http://virtuoso.openlinksw.com/>

⁴⁴<http://sourceforge.net/projects/sesame/>

disponible bajo licencia GPL, es multiplataforma y además ofrece un ENDPOINT SPARQL con la finalidad de consultar los recursos almacenados en el servidor.

3.3.1 Identificación y Extracción de la información

Los datos fueron recolectados de diferentes fuentes, en la actualidad no se cuenta con información sistematizada, completa y confiable esto dificultó la extracción automática de la información sobre las normas, Universidades y Organismos de Educación Superior. Para ello se utilizará los beneficios que brindan las nuevas tecnologías y su forma de organización de la Web Semántica.

Países	Total: 23 países Incluyendo: Argentina, Brasil, Chile, Colombia, Costa Rica, Cuba, Ecuador, El Salvador, Guatemala, Honduras, México, Nicaragua, Panamá, Paraguay, Perú, Puerto Rico, República Dominicana, Uruguay, Venezuela
Tipos de Normas de Educación Superior	Constitución , Leyes, Reglamentos, Resolución, Estatuto, Decretos, Acuerdo
Lenguajes	Español Inglés Portugués Francés
Formatos de archivos	.PDF .DOC

3.3.2 Desarrollo de Vocabularios

Previo a la elaboración del vocabulario se identificó y extrajo la información a representar; para la construcción del vocabulario se reutilizará vocabularios existentes ya que el principio fundamental de Linked Data es la reutilización, con la finalidad si ya existen términos que se acoplen al dominio no sean definidos nuevamente.

De acuerdo a la información a representar sobre Legislación de Educación Superior es posible utilizar los siguientes vocabularios: Dublin Core será utilizado para representar los metadatos de las normas, FOAF para definir a personas u organismos y Dbpedia para obtener información sobre las universidades de Latinoamérica. En vista que en los diferentes vocabularios no se contó con todos los términos necesarios para representar el dominio surgió la necesidad de crear un nuevo vocabulario denominado LUDLAW.

El esquema del vocabulario fue modelado utilizando la herramienta CMAPTOOLS COE, se lo puede apreciar en la figura 1. Los datos que se modelaron corresponden a información clave extraída desde la documentación recolectada.

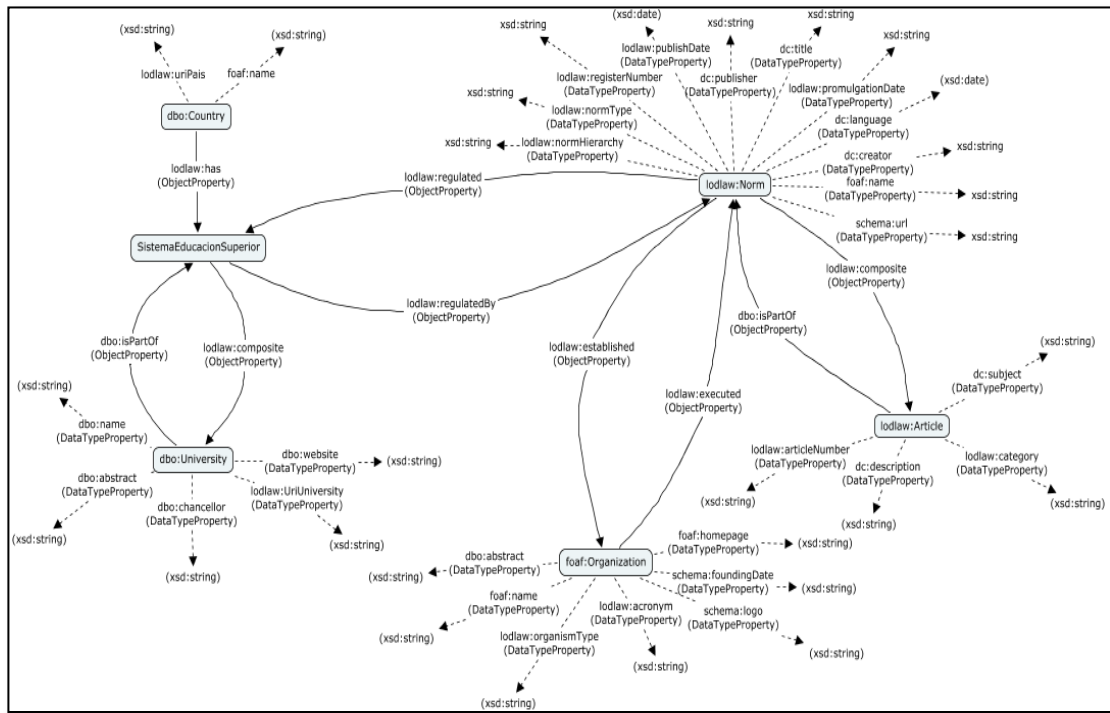


Figura 1: Esquema del vocabulario

3.3.3 Generación de datos RDF

Siguiendo el proceso de publicación de datos el siguiente paso es la generación de datos a formato RDF(W3C, 2008).Para realizar la conversión de datos a RDF se realizó un programa en Java utilizando Jena ya que tiene integrados algunos de los vocabularios a reutilizar como son: FOAF, DUBLIN CORE, RDF, RDFS.

Los vocabularios a reutilizar contienen algunas propiedades que permiten representar la información sobre la Legislación de Educación Superior y para los términos faltantes para cubrir el dominio de la información se creó el vocabulario denominado LUDLAW.

Luego de la conversión obtenemos la información en tripletas o RDF en el siguiente formato:

- Sujeto: una referencia URI, una persona, o un nodo, es el ente al cual nos referimos.
- Predicado: es la propiedad o relación que se desea establecer acerca del sujeto.
- Objeto: valor de la propiedad, o un recurso que puede considerarse el valor definido.

En la Tabla 1 se observa las tripletas o RDF de las normas de educación superior:

Tabla 13 Tripletas de las normas

SUJETO	PREDICADO	OBJETO
http://utpl.edu.ec/educacion/leyes/Norm/NORMA R01	http://utpl.edu.ec/educacion/leyes/ludlaw/normHierarchy	Nivel 3: Legislación Ordinaria
http://utpl.edu.ec/educaci	http://utpl.edu.ec/educacion/leyes/lud	http://www.me.gov.ar/spu/legislacio

on/leyes/Norm/NORMA R01	law/urlNorm	n/Ley_24_521/ley_24_521.html
http://utpl.edu.ec/educacion/leyes/Norm/NORME C01	http://purl.org/dc/elements/1.1/language	Español
http://utpl.edu.ec/educacion/leyes/Norm/NORME C01	http://purl.org/dc/elements/1.1/title	Ley Orgánica de Educación Superior

3.3.4 Publicación de datos

Luego de haber convertido los datos sobre Legislación de Educación Superior a formato RDF, se procede almacenar en un RDF Store en este caso se utilizará Virtuoso, ya que es una herramienta de almacenamiento potente y además cuenta con un SPARQL ENDPOINT para poder consultar la información almacenada en el servidor.

Luego de haber cargado el archivo RDF sobre Legislación de Educación Superior en el motor de almacenamiento de triplas Virtuoso, accedemos al ENDPOINT SPARQL para lo cual colocamos en el navegador lo siguiente: <http://apolo.utpl.edu.ec:8890/sparql>.

¿Cuáles son las normas que en su articulado mencionan sobre el derecho de la Autonomía Universitaria?

```

SELECT ?NORMAS
WHERE
{
?ARTICULO ludlaw:isPartOf ?NORMA.
?NORMA dc:title ?NORMAS
    FILTER REGEX(?ARTICULO, "http://utpl.edu.ec/educacion/leyes/Article/")
    FILTER REGEX(?NORMA, "http://utpl.edu.ec/educacion/leyes/Norm/")
?ARTICULO dc:description ?CATEGORIA
    FILTER REGEX(?CATEGORIA, "AUTONOMIA UNIVERSITARIA")
    FILTER REGEX(?ARTICULO, "http://utpl.edu.ec/educacion/leyes/Article/")
?SISTEMAEDU ludlaw:regulatedBy ?NORMA.    FILTER REGEX(?SISTEMAEDU,
"http://utpl.edu.ec/educacion/leyes/SistemaEducacion/")
?PAISES ludlaw:hasSystem ?SISTEMAEDU.
?PAISES ludlaw:nameCountry ?PAIS
}
GROUP BY ?PAIS

```

NORMAS
Ley de Educación Superior
Constitución de la República De Ecuador
Ley Orgánica de Educación Superior

Ilustración 52 Normas sobre Autonomía Universitaria

3.4 Conclusiones

Después de haber realizado el proyecto de tesis sobre publicación de datos enlazados en el ámbito de legislación de educación superior se concluye lo siguiente:

- Mediante la aplicación de tecnologías y principios de Linked Data se logró cumplir con el objetivo general del tema de tesis “Publicación de datos enlazados en el ámbito de Legislación de Educación Superior” de tal manera que se obtiene datos interconectados y permite reutilizar la información brindando de esta manera valor agregado a la Web.
- Se concluye que la fase de identificación y extracción de datos sobre Legislación de Educación Superior fue la más compleja debido a que la información se encontró dispersa en la web y en diferentes formatos.
- Mediante el uso de estándares como SPARQL, RDF son los más idóneos y se obtuvo mejores resultados al momento de acceder a la información.
- La herramienta Google Refine es muy apta para realizar la limpieza de datos ya que nos permite contar con datos de mejor calidad lo cual asegura brindar respuestas acertadas sobre el dominio de Legislación Universitaria.

3.5 Recomendaciones

- Para realizar la publicación de datos enlazados se recomienda utilizar un proceso de publicación el cual con lleva varios pasos como son: Identificar y seleccionar fuentes de datos, desarrollo de vocabularios, limpieza de datos, generación de datos a RDF, publicación de datos RDF y enlazado de datos RDF con otras fuentes.
- Es aconsejable realizar un análisis previo de los proyectos relacionados sobre el tema de Legislación Universitaria, con la finalidad de mejorar o tomar como referencia para la elaboración e implementación del proyecto.
- Para el elegir el RDF Store se debe realizar un análisis exhaustivo del dominio a representar y se debe considerar los requerimientos antes de proceder a la instalación.

3.6 Bibliografía

1. Berners-Lee, T. (2005). Uniform resource identifier (URI): generic syntax. Obtenido de <http://www.ietf.org/rfc/rfc3986.txt>
2. Bizer, C., Tom Heath, & Tim Berners, L. (s.f.). Linked Data - The Story So Far.
3. Casellas, N. (2012). Legally Linked: Linked Open Data Principles Applied To Code Of Federal Regulations. Obtenido de http://semanticweb.com/legally-linked-linked-open-data-principles-applied-to-code-of-federal-regulations_b29197
4. Casellas, N., Vallbé, J., & Bruce, T. (2011). From Legal Information to Open Legal Data. Recuperado el 25 de 08 de 2012, de <http://dx.doi.org/10.2139/ssrn.1959931>
5. Cifuentes, F. (12 de 08 de 2011). Ontología de Normas para la Biblioteca del Congreso Nacional de Chile.
6. Corcho, O., & Gómez, A. (2010). Mini-curso sobre Linked Data. Recuperado el 25 de 02 de 2012, de <http://www.slideshare.net/ocorcho/linked-data-tutorial-florianpolis>
7. Méndez, E. (24 de 04 de 2012). "RDF: un modelo de metadatos flexible para las bibliotecas digitales del próximo milenio". Obtenido de <http://www.cobdc.org/jornades/7JCD/1.pdf>
8. W3C. (2008). RDF Vocabulary Description Language 1.0: RDF Schema. Recuperado el 06 de 03 de 2013, de <http://www.w3.org/TR/rdf-schema/>