

Integración de datos de los sistemas de información de la Universidad de Cuenca utilizando tecnología semántica

José Zumba-Gómez¹ , Víctor Saquicela² 

¹ Dirección de Tecnologías de Información y Comunicación, Universidad de Cuenca, Av. 12 de Abril s/n y Agustín Cueva, Cuenca, Ecuador 010150.

² Departamento de Ciencias de la Computación, Universidad de Cuenca, Av. 12 de Abril s/n y Agustín Cueva, Cuenca, Ecuador 010150.

Autores para correspondencia: {jose.zumba, victor.saquicela}@ucuenca.edu.ec

Fecha de recepción: 30 de julio de 2017 - Fecha de aceptación: 15 de agosto de 2017

RESUMEN

Hoy en día, la Web Semántica ha alcanzado grandes logros para la integración de la información, ayudando a la gestión y rápida recuperación de información por parte de los usuarios. Por otro lado, en la Universidad de Cuenca se operan diariamente entre 16 y 18 sistemas de información heterogéneos, generando complicaciones a la hora de buscar entidades dentro de todos los sistemas, esto, debido a que no están integrados; por esta razón se propone un modelo de integración utilizando tecnologías de la Web semántica, que permita a los usuarios realizar búsquedas de información más precisas en un contexto global.

Palabras clave: Web semántica, Ontologías, Datos enlazados, NeOn.

ABSTRACT

Today, the Semantic Web has accomplished great progress about data integration, helping the users to manage and recover information quickly. On the other hand, at the University of Cuenca deals daily with about 16 to 18 heterogeneous information systems. The dispersed and not integrated situation of the institution's information systems hinders the task of finding entities. Because of this, we propose an integration model that utilizes semantic Web technologies, which allow users to make more precise findings when searching for information in a global context.

Keywords: Semantic Web, Ontology, Linked Data, NeOn.

1. INTRODUCCIÓN

La Universidad de Cuenca diariamente genera información a través de sus sistemas operacionales, esta es almacenada en distintas bases de datos. Al existir tal cantidad de información y en diferentes fuentes se dificulta realizar búsquedas que respondan con la mayor exactitud a preguntas de distintos usuarios y por tanto no se puede visualizar la información de manera integrada. Hoy en día la Web Semántica juntamente con las ontologías se han convertido en una herramienta útil para lograr una integración de información óptima. Para la Web semántica, la información tiene un significado bien definido, de manera que las personas y las computadoras trabajan de manera cooperativa. Por otro lado, las ontologías son herramientas utilizadas por la Web Semántica para el modelado y gestión del conocimiento. Por los conceptos antes mencionados se ha considerado a la tecnología semántica como la herramienta ideal para recuperar e integrar la información de distintas fuentes de datos. Considerando todos estos aspectos, a través de este trabajo se ha propuesto la elaboración de un modelo ontológico para la representación de los datos y su posterior publicación a través de metodologías de publicación de datos enlazados en el contexto de la Web semántica. La propuesta

consiste en el desarrollo de una solución que permita la construcción de una red ontológica para representar los datos de la estructura organizacional de la Institución y la información académica como, por ejemplo, de unidades académicas y administrativas, facultades, carreras y mallas curriculares, etc. La integración de datos a través de ontologías evitará que los usuarios realicen búsquedas ad-hoc en cada sistema informático para posteriormente consolidar los resultados manualmente.

El resto del trabajo está organizado de la siguiente manera: la sección 2 describe los antecedentes y conceptos de tecnología semántica, así como también análisis de casos exitosos de la aplicación de ontologías y web semántica para la integración de información. La sección 3 detalla el proceso de integración utilizado para este trabajo y finalmente, la sección 4 describe algunas conclusiones y los trabajos futuros.

2. ANTECEDENTES Y TRABAJOS RELACIONADOS

En la WEB existe gran cantidad de información sobre la Web semántica y las ontologías, vistas como una herramienta útil para integrar información registrada en diferentes fuentes de datos. Un ejemplo es el artículo en el que los autores proponen la creación de una ontología para el manejo de datos registrados en diferentes fuentes de la Universidad de la Habana a través de una metodología para llegar a construir un sistema que permite la anotación semántica para la realización de búsquedas federadas y semánticas. Además, hablan de las ventajas para la gestión de la información institucional y que a través de esta aplicación la Institución puede llegar a ser visible en las redes. (Rosell, Senso, & Leiva, 2016). Otro caso revisado fue el de la aplicación de la metodología NeOn para la creación de un modelo ontológico de información académica (planes de curso) y su publicación utilizando tecnología semántica, en este trabajo sus autores concluyen que la tecnología semántica permite la descripción de los datos con un significado claro y preciso, para su posterior publicación con el objetivo de compartir conocimiento y vincular con datos de otras instituciones (Mora & Segarra, 2016).

Para construir una ontología es necesario utilizar una metodología, al respecto existen varias propuestas, es así que en uno de los trabajos revisados se puede ver que sus autores realizan una comparación de diferentes metodologías y métodos para el diseño e implementación de las ontologías; e indica que aunque las metodologías tienen características comunes, estas también se diferencian por la naturaleza de su aplicación y la selección de la metodología adecuada es muy subjetiva, pues depende del alcance de su aplicación. Finalmente, en este trabajo se aplica su metodología a un caso de uso en el dominio de artes plásticas (Guzmán, López, & Durley, 2012). En otro trabajo revisado, los autores indican que para realizar la integración de datos, y publicarlos, se requiere de varios pasos, sin embargo, a su parecer no han sido abarcados todos los necesarios, por esta razón proponen un conjunto de lineamientos metodológicos para realizar las actividades para llevar a cabo todo el proceso y lo prueban en varios contextos gubernamentales como, por ejemplo, GeoLinkedData y AEMETLinkedData (Villazón-Terrazas, Vilches, Corcho, & Gómez, 2011). En otro artículo, al igual que el anterior, también se habla de un método general para publicar linked data y para aplicarlo en diferentes dominios, sin embargo, para sus autores ya no son únicamente 5 pasos, ahora incluyen un paso más llamado Linked (enlaces) (Villazón-Terrazas *et al.*, 2012). Luego de analizar los trabajos relacionados para el desarrollo de este trabajo, la generación, publicación y explotación de los datos enlazados se realizó siguiendo los seis pasos: especificación de requerimientos, modelamiento, generación, establecer enlaces, publicación y explotación (Villazón-Terrazas *et al.*, 2012).

Por otro lado, para la construcción de la ontología se siguió la metodología NeOn, debido a que permite la combinación de escenarios y no es rígida, puesto que sugiere una gran variedad de opciones para construir la ontología (Ramos, Barrera, & Núñez, 2012). Así también, cuando se realiza el modelamiento de una ontología, es recomendable reutilizar, tanto como sea posible, las ontologías y vocabularios existentes, puesto que esta reutilización aumentará la probabilidad de que los datos sean consumidos por aplicaciones que pertenezcan al mismo dominio (Heath & Bizer, 2011).

3. Proceso de integración de datos para los sistemas de información de la Universidad de Cuenca

Para realizar la integración de las fuentes de datos de la Universidad de Cuenca se utilizó como base la metodología propuesta por Villazón-Terrazas *et al.* (2012), que consiste en la aplicación de 6 pasos para la publicación de los datos enlazados. Sin embargo, durante este trabajo se incluyó una fase más que se realiza antes de la explotación, que consistió en realizar la validación de los datos generados para comprobar que la explotación que se realiza a través de un buscador semántico tenga una mayor precisión en los resultados que el usuario busca.

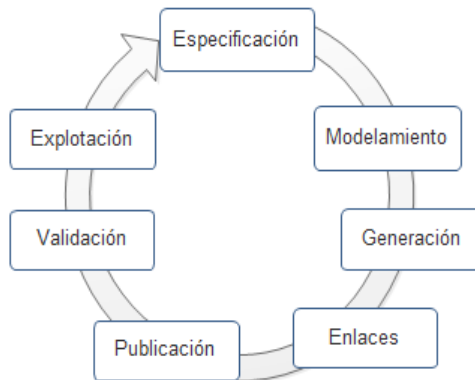


Figura 1. Ciclo de vida para integrar datos de la Universidad de Cuenca.

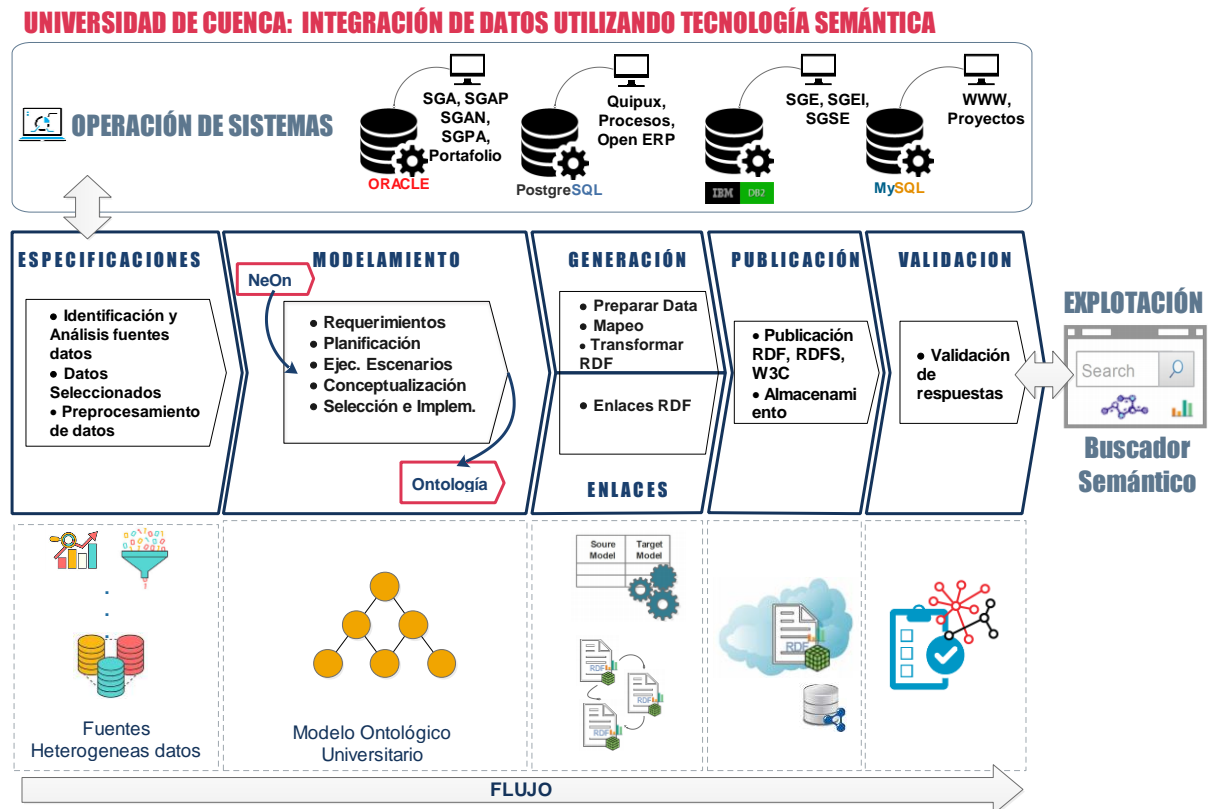


Figura 2. Arquitectura para realizar la integración de datos.

En la Figura 1 se indica gráficamente los pasos del ciclo de vida utilizado en este trabajo para la integración de datos de los diferentes sistemas de información de la Universidad de Cuenca. La arquitectura definida para el desarrollo de este trabajo se puede observar en la Figura 2, en esta se

puede apreciar lo referente a todos los sistemas que se operan en la Universidad de Cuenca y en qué base de datos se almacena la información que genera cada uno de los sistemas. A pesar de que en el gráfico se puede apreciar la mayoría de los sistemas que opera la Universidad de Cuenca, este trabajo está orientado a la publicación de datos enlazados sobre los datos de la estructura organizacional de la Universidad de Cuenca y de los datos académicos de sus carreras y estructura curricular.

3.1. Especificación

Esta etapa consiste en el análisis de los datos que reposan en las distintas bases de datos operadas en la Universidad de Cuenca.

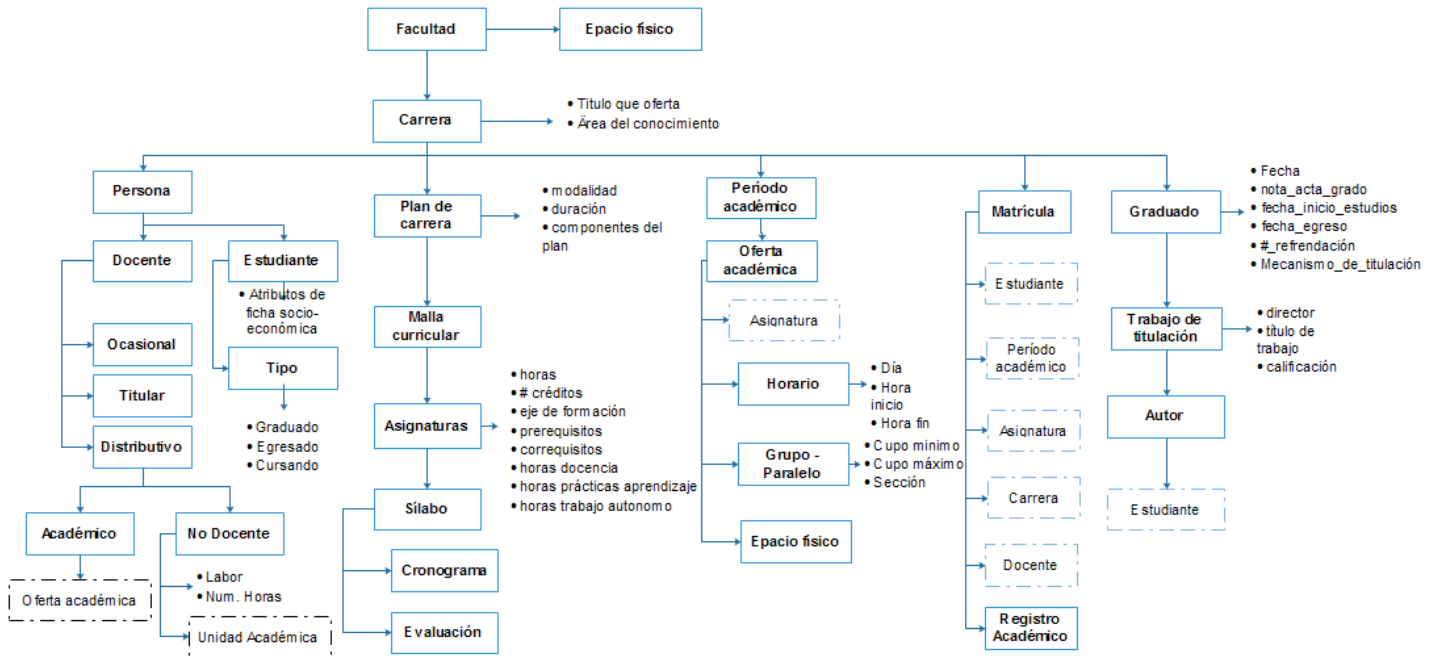


Figura 3. Especificación de requerimientos del SGA.

El análisis permite examinar y seleccionar los datos que serán necesarios y que brinden información con respecto al dominio seleccionado para publicar, para este trabajo se seleccionó la información sobre las dependencias que conforman la estructura organizacional de la Institución, información tomada del estatuto de la Universidad de Cuenca, aprobado en mayo de 2013. La información para el modelamiento del área académica se obtuvo de la documentación del Sistema de gestión académica y de los datos que se almacena en la base de datos Oracle. En la Figura 3 se puede visualizar el análisis de los requerimientos con respecto al sistema de gestión académica.

3.2. Modelamiento

El siguiente paso del ciclo de vida de datos enlazados es el modelamiento de la ontología que permite compartir los datos para que puedan ser procesados por los humanos y por las computadoras. Para realizar el modelamiento de la ontología se aplicó la metodología NeOn, para el caso de la Universidad de Cuenca se seleccionó el Escenario 6: “La reutilización, la fusión y re-ingeniería de recursos ontológicos”. Este escenario fue seleccionado porque en la Web se encontraron algunos recursos ontológicos que pertenecen al mismo dominio que fue modelado para la Universidad, pero que sin embargo debieron ser modificados para que sea de utilidad para el propósito de este trabajo. A continuación, se describen brevemente las principales actividades realizadas para la creación del modelo ontológico en base a la metodología de NeOn. La actividad inicial fue realizar una correcta especificación de requerimientos con el objeto de establecer los requisitos y usos que se dará a la ontología, así como también establecer los usuarios que utilizarán la ontología. En la Tabla 1 se puede ver el resultado de esta actividad, que se trata del documento de especificación de requerimientos

ontológicos (DERO), que permite identificar el conocimiento que se almacenó en la ontología, para facilitar la reutilización de recursos existentes y facilitar la validación de resultados.

Tabla 1. Documento de especificación de requerimientos ontológicos.

Documento de especificación de requerimientos ontológicos	
1.	Propósito
El propósito de construir una ontología universitaria es proveer un modelo genérico de conocimiento que permita describir y presentar conceptos y relaciones existentes en el dominio universitario. El modelo debe ser capaz de responder a consultas referentes a la estructura organizacional y académica de la Universidad de Cuenca. Además, el modelo obtenido debe almacenar la información necesaria para ser utilizado posteriormente en un sistema de búsqueda semántica.	
2.	Alcance
La ontología plantea modelar la estructura organizacional y académica de la Universidad de Cuenca, para lo cual es necesario el almacenamiento de información de diversa índole dentro del contexto universitario. Esta información es planteada en base a:	
<ul style="list-style-type: none"> • Estructura organizacional: Organigrama organizacional aprobado en el estatuto de la Universidad • Estructura académica: Organigrama planteado luego de realizar el estudio de los sistemas y fuentes de datos utilizados en la Universidad 	
3.	Lenguaje de implementación
El lenguaje utilizado para la construcción de la ontología es OWL/XML	
4.	Usuarios finales previstos
<ol style="list-style-type: none"> 1. Autoridades 2. Personal docente y administrativo 3. Estudiantes 	
5.	Usos previstos
<ol style="list-style-type: none"> 1. Consultar información relacionada a la estructura organizacional 2. Consultar información relacionada a la estructura académica 3. Permitir la conexión con otras ontologías 	

Para la especificación de requerimientos funcionales se plantearon preguntas que la ontología debe contestar sobre la estructura organizacional, estructura académica, personas y espacios físicos y recursos bibliográficos. En base a las preguntas establecidas en la especificación de requerimientos ontológicos, se realizó la búsqueda de recursos ontológicos que pertenezcan al dominio de este trabajo, para realizar la búsqueda de ontologías que reúnan los requerimientos identificados se utilizó algunas herramientas online como por ejemplo swoogle¹, Watson², y Sindice³. Además, se realizaron búsquedas manuales en librerías de ontologías como LOV⁴ (Linked Open Vocabularies). Como resultado se obtuvo un conjunto de ontologías candidatas para cada uno de los grupos planteados.

El siguiente paso fue realizar una comparación entre las ontologías para seleccionar la que mejor se adaptó a las necesidades de este trabajo. Para realizar esta comparación se definieron cuatro criterios:

- Propósito similar: El propósito de la ontología debe concordar con el propósito de la nueva ontología.
- Alcance similar: Se verificó que el alcance general o el grupo al cual pertenece sea similar.
- Cobertura de requisitos no funcionales: En este caso se definió que el idioma será el único requisito no funcional a ser considerado.
- Cobertura de requisitos no funcionales: Se comparó el nivel de coincidencia igual o semejante de los términos de la ontología.

¹ <http://swoogle.umbc.edu/2006>

² <http://watson.kmi.open.ac.uk/>

³ <http://sindice.com>

⁴ <http://lov.okfn.org/dataset/lov/>

Una vez definidos los criterios de comparación, se definió también la forma de calificar estos criterios, a través de los siguientes valores cualitativos:

- Si-Totalmente (Si-T): Indica si la ontología cumple de manera exacta con el criterio calificado
- Si-Parcialmente (Si-P): Indica si la ontología candidata cumple de manera parcial con el criterio calificado.
- No (N): Indica que la ontología candidata no cumple con el criterio calificado.
- Desconocido (D): Indica que la ontología candidata no tiene documentación suficiente para determinar si es válida o no para ser reutilizada dentro del criterio calificado.

En la Tabla 2 se puede ver el resultado de aplicar los criterios y su calificación.

Tabla 2. Análisis de recursos ontológicos.

Grupo	Ontología	Propósito	Alcance	Cobertura de requisitos	
				No funcionales	Funcionales
ORGANIZACIÓN	Univ-Bench	Destinado a ser un punto de referencia para la evaluación del rendimiento de los repositorios web semánticos	Ontología de dominio universitario, cubre el área organizacional, personas, publicaciones, trabajos, etc.	N	Si-P
	AISO	Describir la estructura organizacional interna de una institución académica.	Cubre básicamente la estructura organizacional, sin incluir a personas.	N	Si-P
	ORG	Apoyar la publicación de datos vinculados a la información organizacional en varios dominios.	Clasificación de organizaciones y roles, así como extensiones para apoyar información de la organización.	Si	N
	HERO	Ser relevante o al menos conveniente para describir cualquier universidad.	Describe varios aspectos del dominio universitario como la estructura organizativa, personal, roles, ingresos, etc.	N	N
ACADÉMICO	TEACH	Proporciona términos que permiten a los maestros relacionar las cosas en sus cursos.	Describe cursos, asignaturas, materiales, estudiantes, y las relaciones entre ellos	N	Si-P
	LOM	Utilizado para exponer IEEE LOM, un estándar de metadatos para contenidos educativos.	Describe objetos de aprendizaje, requerimientos técnicos, duración, etc.	N	Si-P
	FOAF	Describir personas, sus actividades y relaciones con otras personas y objetos.	Permite describir información básica de las personas, además tiene términos para describir cuentas de internet.	N	Si-P
PERSONAL	VCARD	Describir personas y organizaciones	Permite describir identificación, dirección, contactos, ubicación, propiedades de organización.	N	Si-P

Luego de este proceso se descartaron las ontologías que no cumplieron con los criterios establecidos. Las ontologías seleccionadas fueron las siguientes:

- Estructura organizacional: Univ-Bench
- Estructura académica: HERO
- Personas: FOAF
- Espacios físicos: ROOMS, WGS84 Basic Geo (para latitud y longitud).
- Recursos bibliográficos: BIBO

Luego de analizar todos los recursos y definir cuáles serán reutilizados se procedió a modelar la red ontológica que se puede visualizar en la Figura 4.

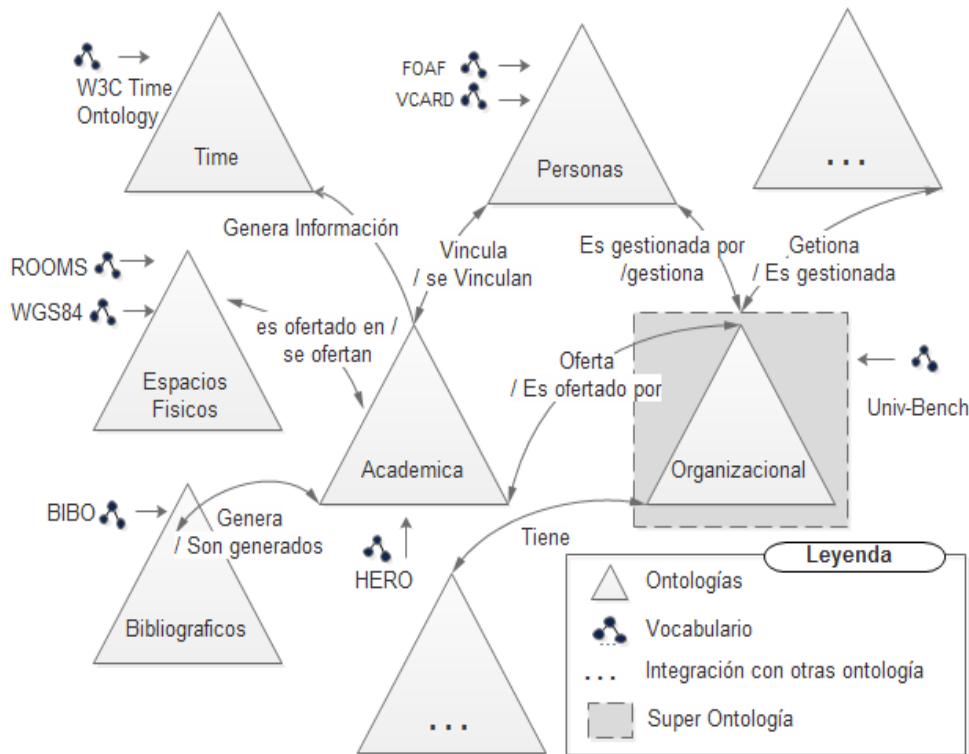


Figura 4. Mapa conceptual de alto nivel de la red ontológica de la Universidad de Cuenca.

A continuación se describe las relaciones existentes en la red ontológica. En primer lugar, se tiene una súper ontología que facilitará la integración con diferentes dominios ontológicos, y que se la denominó *Ontología Organizacional*. Esta súper ontología se vinculó con el dominio *académico*, mediante las carreras; es decir la universidad ofrece carreras y las carreras son ofertadas por las unidades académicas; la universidad es gestionada por *Personas*; las *Personas* gestionan la Universidad; a las carreras se vinculan estudiantes o docentes que son *Personas*; y los estudiantes o docentes cursan o dictan clases en los cursos ofertados. Los cursos se ofertan en *espacios físicos*. Los cursos se dictan en ciertos *periodos de tiempo*. Los cursos generan recursos bibliográficos, como por ejemplo trabajos de titulación, publicaciones, etc.

La ontología fue desarrollada utilizando el lenguaje ontológico OWL (Web Ontology Language), a través de la herramienta Protégé⁵.

3.3. Generación

El siguiente paso del ciclo de vida de datos enlazados es la conversión de los datos a formato RDF (Resource Description Framework), para la generación de los grafos se realizó primero un proceso de limpieza de los datos de manera que estos se encuentren consistentes y libre de errores.

En primer lugar, se hizo una limpieza de los datos correspondientes a la estructura organizacional, puesto que en la base de datos en donde se almacena esta información existen dependencias duplicadas, por lo que se hizo un mapeo para recuperar las dependencias sin duplicidades, de manera que no se den futuros errores en el poblado de la ontología. Otro proceso de limpieza se realizó sobre la información correspondiente a las carreras de la Universidad, debido a que en las bases de datos existe información sobre carreras auxiliares para que el sistema de gestión académica cumpla con ciertas funcionalidades. Por esta razón, se procedió a seleccionar únicamente las carreras válidas y sin duplicidad, realizando, de igual manera que en el caso anterior, un mapeo en caso de que una carrera tenga más de un registro con el mismo nombre. Estos procesos de limpieza se realizaron con la finalidad de llegar a una publicación exitosa que permita explotar la información con un alto grado de

⁵ <https://protege.stanford.edu/>

confianza. Para realizar la limpieza de los datos, se modelaron y ejecutaron procesos ETL a través de la herramienta Pentaho data integration⁶. Una vez que los datos fueron procesados, consistentes y sin errores, estos se almacenan en una base de datos. El siguiente paso fue realizar la descripción semántica utilizando los vocabularios de las ontologías que fueron modeladas en la fase anterior. Para cargar los vocabularios de las ontologías se utilizó la plataforma de generación de Linked Data a través de la plataforma Pentaho⁷. A continuación se procedió a realizar los mapeos necesarios de los datos registrados en la base de datos intermedia con las ontologías definidas. La última actividad realizada en esta fase fue la transformación de los datos al formato RDF a través de la plataforma, por esta razón se indica que la anotación semántica en este trabajo fue un proceso semi-automático. Finalmente, los grafos generados mediante la plataforma fueron almacenados en una base de datos de tripletas⁸, que servirá como base para las consultas en el momento de la explotación. En la Figura 5 se puede ver gráficamente las herramientas utilizadas durante la fase de generación.

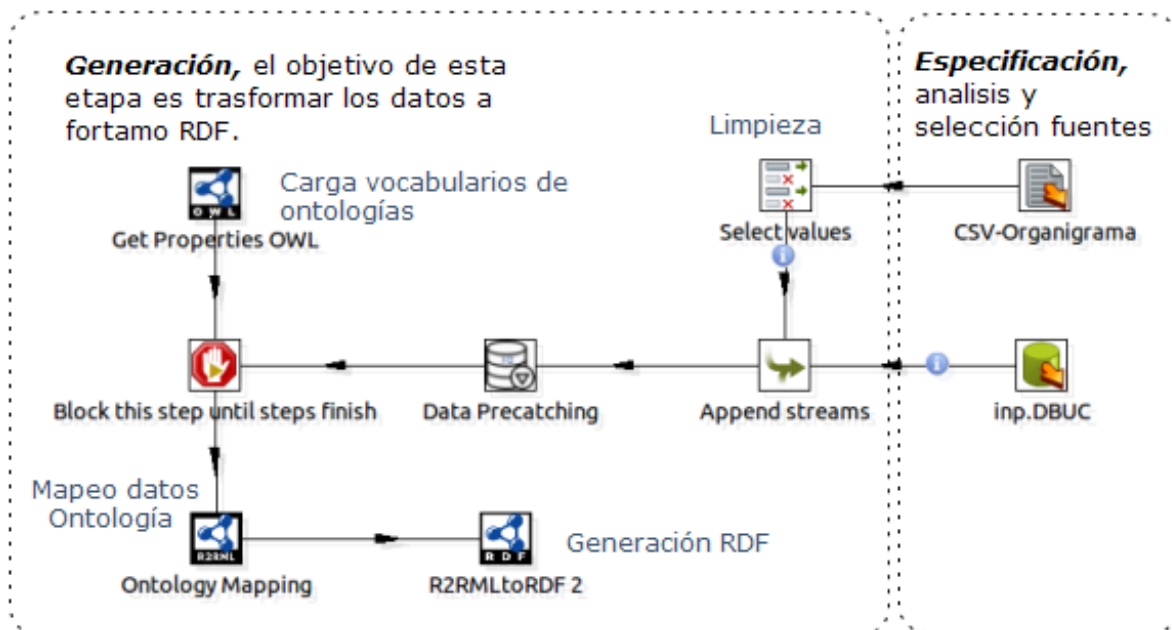


Figura 5. Herramientas utilizadas en la fase de generación.

3.4. Enlaces

Al momento de poblar la ontología se fueron generando algunos enlaces internos con otras fuentes de datos con los que cuenta la Universidad, debido a que se utiliza un modelo ontológico común, permitiendo a los usuarios abrir el abanico de información a consultar. Así también la definición de los enlaces entre los RDFs que han resultado de la fase de generación con grafos similares existentes en la Web, es decir con fuentes de datos externas, es una tarea bastante importante, puesto que pueden aportar de manera significativa a los recursos publicados como datos enlazados en la Web; enriqueciendo los recursos de los que se dispone y permitiendo que los resultados obtenidos tengan mayor alcance. El establecimiento de enlaces con las fuentes externas se propone como un trabajo futuro a este trabajo. Para realizar esta actividad: primero se deberán definir los recursos que podrán ser enlazados y se debe también identificar las posibles fuentes de datos externas que pertenezcan al mismo dominio.

⁶ <http://www.pentaho.com/>

⁷ <https://ucuenca.github.io/lodplatform/>

⁸ <https://jena.apache.org/documentation/fuseki2/>

3.5. Publicación

En esta etapa del ciclo de vida se visibiliza la información obtenida en la fase de generación de manera que los grafos en RDF puedan ser accedidos a través de la Web. Esto se logra mediante las siguientes tareas: (1) Publicación y almacenamiento de las fuentes de datos transformadas en tripletas RDF, (2) Publicación de los metadatos de las fuentes de datos antes mencionados; y (3) permitir que los datos se encuentren actualizados para su posterior recuperación (Villazón-Terrazas *et al.*, 2011).

Para la publicación de las tripletas obtenidas en la generación, se requiere un servidor de Linked Data triplestore, que guardará y permitirá recuperar RDF. El triplestore a utilizar es Apache Marmotta⁹ que se encuentra instalado en el centro de datos de la Universidad de Cuenca.

3.6. Validación

Una vez que se cumplió la fase de publicación, se vio la necesidad de incluir, en el ciclo de vida, una fase de validación. Esta fase ha sido incluida con el objeto de asegurar que la ontología fue modelada correctamente y que refleja la realidad para la que fue concebida. Durante la actividad de validación se comprueba que, durante la fase de modelamiento, se utilizaron los conceptos correctos y se evalúan los posibles errores de sintaxis cometidos durante la construcción misma de la ontología (Barrera, Nuñez, & Ramos, 2012).

The screenshot shows the Fuseki Query interface. On the left, the SPARQL Query is displayed: `PREFIX uy: <http://swat.cse.lehigh.edu/onto/univ-bench.owl#> PREFIX hero: <http://www.UniversityReferenceOntology.org/HERO#> PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> PREFIX core: <http://vivoweb.org/ontology/core#> PREFIX as: <http://purl.org/vocab/aissq/schema#> select ?nombre_facultad where { ?facultad rdf:type as:Faculty. ?facultad uy:name ?nombre_facultad } ORDER BY ASC(?nombre_facultad)`. The output is set to 'Text'. On the right, the results are displayed in a table with the column 'nombre_facultad' and the following rows: "FACULTAD DE ARQUITECTURA Y URBANISMO", "FACULTAD DE ARTES", "FACULTAD DE CIENCIAS AGROPECUARIAS", "FACULTAD DE CIENCIAS DE LA HOSPITALIDAD", "FACULTAD DE CIENCIAS ECONÓMICAS Y ADMINISTRATIVAS", "FACULTAD DE CIENCIAS MEDICAS", "FACULTAD DE CIENCIAS QUIMICAS", "FACULTAD DE FILOSOFIA, LETRAS Y CIENCIAS DE LA EDUCACION", "FACULTAD DE INGENIERIA", "FACULTAD DE JURISPRUDENCIA Y CIENCIAS POLITICAS Y SOCIALES", "FACULTAD DE ODONTOLOGIA", and "FACULTAD DE PSICOLOGIA".

Figura 6. Respuesta a pregunta planteada en el documento de requerimientos.

La forma de validación que se aplicó fue plantearse preguntas y comprobar que el conjunto de datos RDF puedan responder a estas. Por otro lado, se realizaron consultas a las bases de datos por los expertos de los sistemas informáticos para obtener la información que conteste a las preguntas planteadas. Por lo que, se comprobó que la ontología fue modelada correctamente, puesto que pudo contestar a las preguntas planteadas en el documento de requerimientos. Como ejemplo, se puede ver en la Figura 6 la respuesta a una de las preguntas planteadas: “Cuáles son las facultades de la Universidad de Cuenca? Para obtener las respuestas a las preguntas, se utilizó el lenguaje de consultas SPARQL.

3.7. Explotación

En el ciclo de vida de los datos enlazados, la última fase es la explotación. No es suficiente que los datos estén publicados en la Web, por esta razón, como parte de este trabajo se construyó un prototipo de un buscador semántico para que los datos publicados puedan ser consumidos. Todas las consultas que se realizan a través de este buscador semántico fueron implementadas a través de consultas escritas en lenguaje SPARQL.

⁹ <http://marmotta.apache.org/>

4. CONCLUSIONES Y TRABAJOS FUTUROS

El modelo ontológico construido ha permitido representar el dominio de la estructura organizacional de la Universidad de Cuenca, así como también todo lo referente al área académica, de manera que su publicación sea de gran utilidad para toda la comunidad universitaria, puesto que los usuarios pueden consultar sobre datos importantes de la Universidad, sus carreras, oferta académica y datos importantes como las calificaciones y datos socioeconómicos de sus estudiantes.

Como se mencionó anteriormente, uno de los trabajos futuros que se plantea es la definición de enlaces con fuentes de datos externas, de manera que se pueda enriquecer los datos originales obtenidos en este trabajo. Se espera que este trabajo sirva de guía para futuros desarrollos de integración de datos, para la Universidad de Cuenca, como por ejemplo información sobre la evaluación docente o sobre las investigaciones que los docentes realizan como parte de la Institución.

AGRADECIMIENTO

Este trabajo fue realizado como parte de los requisitos previo a la obtención del título “Máster en Gestión Estratégica de Tecnologías de la Información” impartido por la Facultad de Ingeniería de la Universidad de Cuenca.

REFERENCIAS BIBLIOGRÁFICAS

- Barrera, M., Nuñez, H., Ramos, E. (2012). *Ingeniería ontológica*. Lecturas en Ciencias de la Computación. Universidad Central de Venezuela Facultad de Ciencias, Escuela de Computación. 63 p.
- Bizer, C. (s.f.). *D2RQ Accessing Relational Databases as Virtual RDF Graphs*. Obtenido de d2rq.org
- Guzmán, J., López, M., Durley, I. (2012). Metodologías y métodos para la construcción de ontologías. *Scientia et Technica*, XVII(50), 133-140.
- Heath, T., Bizer, C. (2011). *Linked data evolving the web into a global data space*. San Rafael, CA: Morgan & Claypool Publishers. <https://doi.org/10.2200/S00334ED1V01Y201102WBE001>
- Mora, M. B., Segarra, V. (2016). Modelo ontológico para la representación de datos académicos y su publicación con tecnología semántica. *Opción*, 32(10), 267-282.
- Rosell, Y., Senso, J., Leiva, A. (2016). Diseño de una ontología para la gestión de datos heterogéneos en universidades: marco metodológico. *Revista Cubana de Información en Ciencias de la Salud*, 27(4), 545-567.
- Villazón-Terrazas, B., Vila-Suero, D., Garijo, D., Vilches, L., Poveda, M., Mora, J., . . . Gómez, A. (2012). *Publishing linked data - There is no one-size-fits-all formula*. 9 p. Disponible en <http://ceur-ws.org/Vol-877/poster5.pdf>
- Villazón-Terrazas, B., Vilches, L., Corcho, O., Gómez, A. (2011). *Methodological guidelines for publishing government linked data*. In: Wood, D. (Ed.). *Linking Government Data*. Springer, New York, NY.