

Un enfoque de Linked Data para garantizar la interoperabilidad semántica e integridad de datos académicos universitarios

Efren Narvaez¹ and Nelson Piedra¹

Universidad Técnica Particular de Loja, Loja 1101608, Ecuador
{ernarvaez,nopiedra}@utpl.edu.ec

Abstract. The growing amount of information in the organizations has led to problems of interoperability and data integrity, thereby forcing development and adoption of new technologies and approaches to give solution to these problems. A solution to this problem is to use the approach that provides the semantic web to achieve interoperability between data from different sources and information systems. Other one of the advantages that offers the semantic web is the construction of descriptive models of information based on ontologies like more complete structures for the description of data. This allows a formal description of them, in addition to their semantic and syntactic representation. This paper describes the design of a RESTful Web services-based architecture that guarantees the semantic interoperability and integrity of academic data of Loja's Particular Technical University. The implementation of the solution proposed in this work is based on the development of a prototype of a RESTful API which provides general information about students, teachers, authorities and academic components. These data are stored in Apache Marmotta's triple store, an open implementation of a linked data platform.

Keywords: Semantic Web, Data, RDF, Ontology, RESTful, Integration, Interoperability

Resumen La creciente cantidad de información en las organizaciones ha generado problemas de interoperabilidad e integridad de datos, obligando así al desarrollo y adopción de nuevos enfoques y tecnologías que permitan dar solución a estos problemas. Una solución para este problema es emplear el enfoque que brinda la web semántica para lograr la interoperabilidad entre datos de diferentes fuentes y sistemas de información. Otra de las ventajas que ofrece la web semántica es la construcción de modelos descriptivos de datos basados en ontologías como estructuras más completas para la descripción de datos. Esto permite una descripción formal de los mismos, además de su representación semántica y sintáctica. El presente trabajo describe el diseño de una arquitectura basada en RESTful web services que garantice la interoperabilidad semántica

e integridad de datos académicos de la Universidad Técnica Particular de Loja. La implementación de la solución propuesta en este trabajo se basa en el desarrollo de un prototipo de un RESTful API el cual provee información general acerca de Estudiantes, Docentes, Autoridades y Componentes Académicos. Estos datos se encuentran almacenados en el triple store de Apache Marmotta, una implementación abierta de una plataforma de datos enlazados.

Palabras claves: Web Semántica, Datos, RDF, Ontología, Integración, Interoperabilidad

1. Introducción

Actualmente las organizaciones poseen sistemas informáticos que diariamente alimentan a grandes bases de datos y que con el tiempo estas adquieren grandes volúmenes. Estas bases de datos con grandes volúmenes de información tienden a convertirse en un gran problema para las organizaciones que no poseen una arquitectura de aplicaciones y de datos correctamente definida [17]. El manejo de grandes volúmenes de datos alimentados por varios sistemas informáticos conlleva a problemas de redundancia e integridad en los datos, razón por la que se dificultan las tareas de comunicación e intercambio de información entre aplicaciones [2].

El problema que actualmente le impide a la Universidad Técnica Particular de Loja (UTPL) lograr una interoperabilidad de datos entre las diferentes aplicaciones que posee radica en la heterogeneidad de las fuentes de información que posee, los grandes volúmenes de datos, la existencia de entidades ambiguas y poco consistentes que estas aplicaciones generan.

La interoperabilidad semántica de datos entre aplicaciones se presenta como una solución viable a este problema, definiendo un estándar de tecnologías, formatos de datos, reglas y medios de comunicación entre las aplicaciones [7]. Es necesario adoptar tecnologías de web semántica para reducir problemas de integridad de datos.[8].

En el contexto de educación superior, las universidades y organizaciones similares almacenan la información que producen en diferentes sistemas de información. Sin embargo, la falta de uso de estándares abiertos, la aplicación de un enfoque de calidad de datos, y la adopción de una aproximación semántica han causado problemas a la hora de integrar y re-usar contenidos a través de la Web [4] [13]. En este sentido, tanto del mundo académico como de la industria, están comprometidos con diferentes esfuerzos para describir semánticamente datos e información [18], mejorar la calidad de datos que se comparten [18], y crear aplicaciones basadas en grafos de conocimiento (knowledge graph, KG) que promuevan la explotación de datos con semántica explícita [5]. A la fecha Hay importantes avances en la aplicación de la Web Semántica en el contexto educativo. En [11], [14] y [15] se proponen arquitecturas de integración e interoperabilidad distribuida de repositorios digitales, basadas en un enfoque de Web Semántica y tecnologías de Linked Data. Estas arquitecturas han sido probadas

con éxito en tareas de integración de repositorios institucionales pertenecientes a universidades, y han habilitado un potencial entorno de consultas federadas dentro de este contexto.

En este trabajo se propone un marco de interoperabilidad semántica que permita integrar la información general de Docentes, Estudiantes, Componentes Académicos y la estructura organizacional de la UTPL (Gobierno UTPL). Los datos y su semántica son descritos usando el modelo RDF [6], esto en formato de tripletas los cuales son almacenados dentro de un triplestore y expuestos a través de la plataforma de linked data (LDP) como lo es Apache Marmotta. Se cubren los procesos de extracción, transformación y carga que alimente de datos a esta plataforma [9]. Además, se incluye el desarrollo de un API RESTful que permita la interoperabilidad de esta plataforma con las distintas aplicaciones que necesiten acceder a estos datos.

El piloto descrito en este trabajo usa información del Sistema de Información Estratégica Centralizada (SIEC-UTPL) la cual maneja una copia de la información de varios sistemas de la UTPL, entre otros:

- El Entorno Virtual de Aprendizaje - (EVA).
- El Nuevo Sistema de Gestión Académica - (NSGA).
- Distributivo.
- El Sistema de Infracción Académica Científica - (SIAC).
- FileMaker (Base de datos de recursos humanos).
- El Sistema de Información Estratégica Centralizada - (SIEC).
- El Sistema de Exalumnos.

2. Metodología

Conforme a las recomendaciones establecidas por la W3C se debe seguir un ciclo de publicación de datos enlazados. El ciclo de vida usado en esta investigación se basa en el modelo propuesto por [12] el mismo que se encuentra dentro del dominio de datos relacionados con la educación superior. La información a ser extraída, depurada, transformada en datos enlazados y cargada en Apache Marmotta se ubica dentro de este dominio.

2.1. El problema

Gran volumen de datos, diversidad de tecnologías y esquemas de almacenamiento, dispersión y ambigüedad, se evidencia en las fuentes de datos que se analizan en este trabajo. La duplicidad de registros y entidades existentes en cada una de las diferentes bases de datos Institucionales analizadas (Financiero, Académico, Recursos Humanos, Biblioteca, Editorial, Unidad de Graduados.), evidencia claramente las dificultades técnicas y logísticas para la integración de estas fuentes y sistemas de información.

Diseño de la solución La solución planteada consta de dos etapas.

La primera: diseñar un modelo ontológico de datos en base al dominio definido, que permita la representación semántica de los datos y la interoperabilidad semántica entre fuentes de datos heterogéneas y dispersas. El grafo de conocimiento resultante se publicará en un triplestore basado en Apache Marmotta.

La información a ser extraída, transformada y cargada en Apache Marmotta será acerca de las siguientes entidades:

- Docentes.
- Autoridades UTPL.
- Estudiantes.
- Componentes Educativos

La segunda: diseñar e implementar una arquitectura de software robusta de alta disponibilidad que garantice la interoperabilidad semántica e integridad de datos entre sistemas de información. Para este diseño se utiliza el modelo de vistas 4+1 propuesto por [10] que sigue el estándar “IEEE 1471-2000” (Recommended Practice for Architecture Description of Software-Intensive Systems) [1]. Utilizado para describir la arquitectura de un sistema software basado en el uso de múltiples vistas.

2.2. Implementación de la solución.

Conforme a las recomendaciones establecidas por la W3C se debe seguir un ciclo de publicación de datos enlazados. La opción seleccionada es el ciclo de vida propuesto por [16] el mismo que se encuentra dentro del dominio de datos relacionados con educación superior; este ciclo propone un marco de trabajo para la publicación de datos enlazados, que inicia con una fase de observación del modelo de negocio, identificación y selección de las fuentes de datos, modelamiento, generación de datos RDF, publicación, consumo y visualización.

La solución propuesta se basa en las siguientes tecnologías y estándares.

- **Almacenamiento:** Apache Marmotta kiwi triplestore.
- **Lenguajes de modelado de datos:** RDFS y OWL.
- **Descripción de recursos:** RDF.
- **Lenguaje de consulta:** SPARQL.
- **Protocolo de acceso a datos:** HTTP.
- **Arquitectura de Servicios:** RESTful.

A continuación, se describe el proceso desarrollado:

Identificación de fuentes. Las fuentes de datos identificadas para obtener información son las siguientes:

- El Nuevo Sistema de Gestión Académica - (NSGA).
- FileMaker (Base de datos de recursos humanos).
- El Sistema de Información Estratégica Centralizada - (SIEC).
- El Sistema de Exalumnos.

Modelamiento del vocabulario. Basados en el marco de trabajo propuesto por [16] y [4] se realiza mapeo entre conceptos y entidades de almacenamiento físico, de tal forma que se pueda seleccionar y reusar la mayor cantidad de vocabularios ya existentes y estandarizados [3]. Que permitan modelar la ontología sobre la cual van a ser almacenados los datos. En este trabajo se presenta el uso de vocabularios semánticos que enriquecen los metadatos con a una descripción semántica basada en RDF. Estos conceptos son accedidos a través de un URI o identificador único el cual se conecta con otro grafo RDF que contiene toda la información detallada de lo que ese dato representa en el mundo real y las relaciones que este tiene con otros recursos. De esta forma el dato adquiere una representación más amplia y clara al unirse a la nueve de datos enlazados que existen en la web actual.

Transformación de datos a RDF. Luego de la extracción, depuración y modelado de la ontología es posible definir un proceso de transformación de datos relacionales a un esquema RDF que se almacene dentro de un triple store, en este caso el triple store de Apache Marmotta. En este caso particular se diseñó este proceso con la ayuda de una aplicación desarrollada en Java utilizando librerías como: RDF4J, Apache Marmotta API, JDBC y SESAME un framework para linked data.

Publicación de Datos. Uno de los puntos finales es la publicación de la data transformada a RDF, esta data reposa sobre una plataforma para Linked data como lo es Apache Marmotta. La publicación se la realiza justo después de que los datos fueron convertidos a RDF con la ayuda del API de Apache Marmotta se envía a través del protocolo HTTP al servidor en donde se encuentra el triple store de Marmotta.

Consumo y Visualización. Los datos pueden ser accedidos de dos formas diferentes. La prima a través del RESTful API construido sobre la arquitectura propuesta como parte de la solución y que retorna datos en cuatro diferentes formatos: JSON, RDF, XML, JSON-LD. La segunda a través del SPARQL endpoint que Apache Marmotta proporciona en su consola de administración.

2.3. Desarrollo de la aproximación.

A partir de la solución propuesta se identificaron las fuentes de datos que contiene información sobre las entidades seleccionadas. El cuadro 1 a continuación muestra un mapeo de las entidades y los sistemas y bases de datos en las que se encuentra almacenada información relacionada.

Cuadro 1. Relación Entre las Fuentes y las Entidades

Sistema	Docente	Autoridad	Componente Educativo	Estudiante
Nómina	x	x		
RRHH	x	x		x
SIAC	x			
SIEC	x		x	x
Distributivo	x		x	x
NSGA	x		x	x

Una vez identificadas las fuentes de información se define y/o selecciona los vocabularios a ser utilizados en cada una de las propiedades de las entidades.

El cuadro 2 referente a la entidad Persona muestra una relación entre las fuentes de datos, sus atributos y la equivalencia de estos atributos en un vocabulario RDF.

Cuadro 2. Mapeo entre atributos que describen a personas y propiedades en RDF.

Personas			
Fuente de Datos		Equivalencia Vocabularios RDF	
Sistema	Tabla	Atributo	Vocabulario:Propiedad
SIEC	TDI_ENTIDAD	Nombres	Foaf:name
		Apellidos	foaf:FamilyName
		Cédula	dcterms:identifier
		Edad	Foaf:age
		Fecha Nacimiento	foaf:birthday
		Estado Civil	lodutpl:MaritalStatus
		Lugar de Nacimiento	schema:birthPlace
		Género	foaf:gender
		Email	foaf:mbox
		Teléfono	foaf:phone

La entidad persona abarca toda la información general que una entidad de este tipo posee, pero es necesario definir atributos propios de las subentidades como Estudiante, Docente, Autoridad UTPL. De tal forma que se pueda realizar una presentación semántica de cada una de las entidades.

A continuación, el cuadro 3 muestra información de Estudiante y las relaciones de sus fuentes, atributos y equivalencias con los diferentes vocabularios RDF.

Cuadro 3. Mapeo entre atributos que describen a estudiantes y propiedades en RDF

Estudiante			
Fuente de Datos		Equivalencia Vocabularios RDF	
Sistema	Tabla	Atributo	Vocabulario:Propiedad
SIEC	TDI_ESTUDIANTE	Fecha Inicio Estudios UTPL	schema:startDate
		Fecha Fin Estudios	schema:endDate
		Estado Académico	lodutpl:AcademicStatus
		Nivel Académico	vivo:AcademicDegree

El cuadro 4 muestra la información específica de la subentidad Docente.

Cuadro 4. Mapeo entre atributos que describen a docentes y propiedades en RDF.

Docente			
Fuente de Datos		Equivalencia Vocabularios RDF	
Sistema	Tabla	Atributo	Vocabulario:Propiedad
SIEC	TDI_DOCENTE	Fecha Inicio Labores UTPL	schema:startDate
		Fecha Fin Labores UTPL	schema:endDate
		Estado Laboral	lodutpl:LaboralStatus
		Nivel Académico	vivo:AcademicDegree
		Extensión Telefónica	mads:extension

Cuadro 5. Atributos de la Subentidad Autoridad UTPL

Autoridad UTPL			
Fuente de Datos		Equivalencia Vocabularios RDF	
Sistema	Tabla	Atributo	Vocabulario:Propiedad
SIEC	TDI_CARGODOCENTE TDI_CARGOPERFIL TDI_CARGO TDI_PERFIL	Fecha Inicio Labores UTPL	schema:startDate
		Fecha Fin Labores UTPL	schema:endDate
		Estado Laboral	lodutpl:LaboralStatus
		Nivel Académico	vivo:AcademicDegree
		Extensión Telefónica Cargo	mads:extension org:role

El cuadro 6 a continuación muestra la información de la entidad componente académico.

Cuadro 6. Atributos y Equivalencia de la Entidad Componente Educativo

Componente Académico			
Fuente de Datos		Equivalencia Vocabularios RDF	
Sistema	Tabla	Atributo	Vocabulario:Propiedad
SIEC	TDI_COMP_ACADMC	Código del Componente	aiis:Code
		Nombre del Componente	aiiso:name
		Descripción del Componente	aiiso:Description
		Vigencia del Componente	lodutpl:ValidityStatus
		Ciclo Correspondiente del Componente	aiiso:Module
		Número de Créditos del Componente	lodutpl:AcademicCredit
		Departamento Responsable del Componente	aiiso:Department
		Titulación que Oferta el Componente	aiiso:Faculty
	Tipo de Componente	lodutpl:SubjectType	

Posterior a la identificación de los vocabularios, propiedades, rangos y dominios se modela la ontología sobre la cual se van a almacenar los datos que posteriormente se transformaran a un formato de tripletas RDF.

La figura 1, muestra el modelo ontológico diseñado para la entidad componente académico.

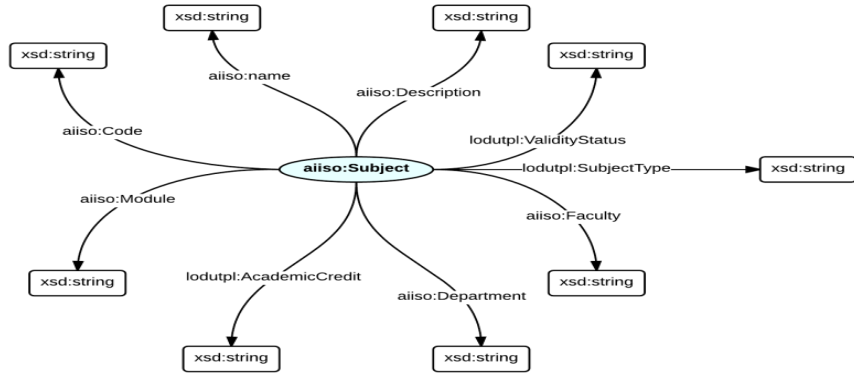


Figura 1. Grafo Utilizado Para Describir un Componente Académico.

La figura 2, muestra el modelo ontológico para la entidad persona:

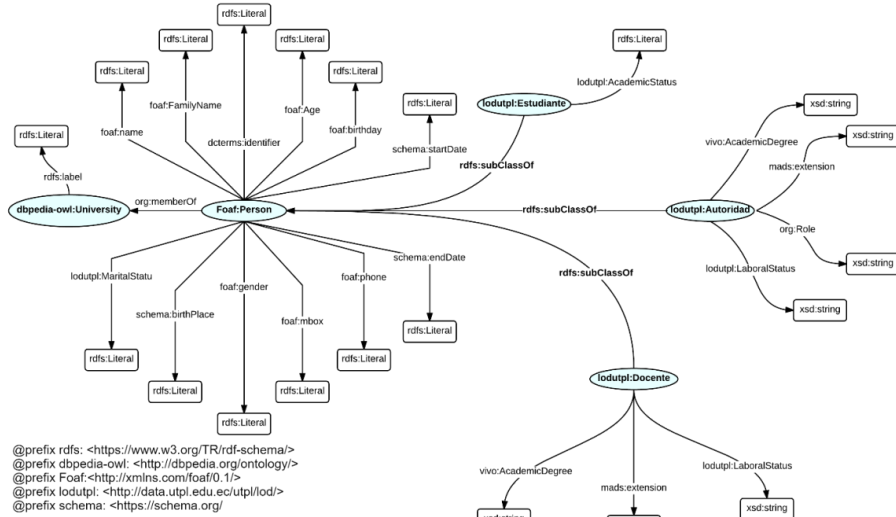


Figura 2. Grafo Utilizado Para Describir a la entidad persona y sus subentidades.

2.4. Transformación de datos a RDF y publicación de datos.

Para realizar la transformación de los datos a una estructura semántica se utiliza tecnología Java. El framework SESAME permite armar una descripción

de un recurso en RDF para su posterior serialización en los diferentes formatos que la W3C determina para Linked Data.

Al finalizar toda la transformación de la data se procede a cargar el RDF sobre el triplestore de Apache Marmotta la plataforma para Linked Data que ésta investigación selecciona como repositorio de datos. El cuadro 9 muestra los resultados producto de la generación y carga de un total de 271.140.00 tripletas que corresponden a los datos de las cuatro entidades seleccionadas en el caso de aplicación.

Cuadro 7. Tiempos y recursos, transformación y carga.

Resumen Transformación y Carga				
Proceso	Número de Tripletas	Formato de Serialización	Memoria RAM Ocupada	Tiempo de Ejecución.
Transformación y Carga	271.140.00	application/x-turtle	7 MB	1:88(m:ss)

Apache Marmotta realiza una negociación de contenido básica en la cual brinda la opción de presentar los datos como tripletas en una página HTML o en un formato RDF. Esta negociación de contenido depende de cómo se acceda al recurso. La figura 3 muestra la visualización de la información de un recurso en forma de una página web puesto que se accede al URI de dicho recurso desde un navegador.

The screenshot shows the Apache Marmotta web interface. At the top left is the logo for UTPL Marmotta. The main content area is titled 'Triples' and displays a table of properties and values for the resource 'http://gestiondatos.utpl.edu.ec:8090/marmotta/resource/MAURO_RODRIGO'. The table has four columns: 'property', 'has value', 'context', and 'info'. The properties listed include 'lodutpl:laboralStatus', 'lodutpl:maritalStatus', 'schema:endDate', 'schema:nationality', 'schema:startDate', 'rdf:type', 'rdfs:subClassOf', 'foaf:birthDay', 'foaf:familyName', 'foaf:gender', 'foaf:mbox', 'foaf:name', and 'foaf:phone'. The values are mostly strings, and the context is 'context:person' for most entries.

property	has value	context	info
lodutpl:laboralStatus	"A"^^xsd:string	context:person	
lodutpl:maritalStatus	"CASADO"^^xsd:string	context:person	
schema:endDate	"2009-01-24 00:00:00.0"^^xsd:string	context:person	
schema:nationality	"ECUATORIANO"^^xsd:string	context:person	
schema:startDate	"2000-01-24 00:00:00.0"^^xsd:string	context:person	
rdf:type	lodutpl:Docente	context:person	
rdfs:subClassOf	foaf:Person	context:person	
foaf:birthDay	"1967-01-24 00:00:00.0"^^xsd:string	context:person	
foaf:familyName	"AVILES SALVADOR"^^xsd:string	context:person	
foaf:gender	"M"^^xsd:string	context:person	
foaf:mbox	"mraviles@utpl.edu.ec"^^xsd:string	context:person	
foaf:name	"MAURO RODRIGO"^^xsd:string	context:person	
foaf:phone	"1967-01-24 00:00:00.0"^^xsd:string	context:person	

UTPL LOD es un proyectos de datos abiertos de la Universidad Tecnica Particular de Loja (Ecuador)

Figura 3. Visualización de la infracción del recurso Mauro Rodrigo.

3. Explotación de datos.

El acceso a la información se la puede realizar de dos formas diferentes, la primera desde un RESTful API y la segunda directamente del SPARQL endpoint de Apache Marmotta. Cada una de estas opciones con tiempos de respuesta diferentes. El RESTful API semántico que se construyó como parte de la solución fue comparado con el API-UGTI que es el que la UTPL ocupa actualmente en varios de los sistemas que posee, Se consideraron características como los tiempos de respuesta promedio entre la API semántico, el API-UGTI y una conexión directa a base de datos desde una aplicación por medio de un JDBC.

El resultado de esta comparativa se lo puede visualizar en la figura 4 a continuación.

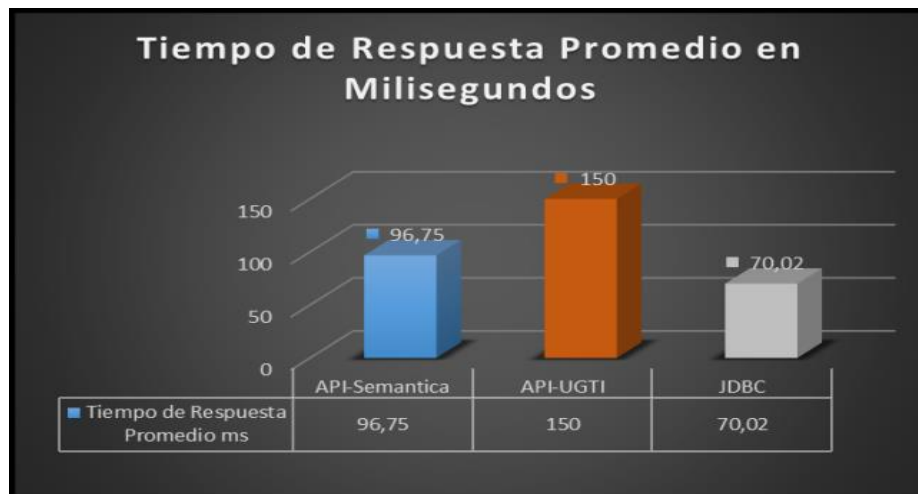


Figura 4. Comparativa de tiempos entre la solución y otras alternativas.

Cabe aclarar que el mejor tiempo lo tiene la conexión por medio de un JDBC con una diferencia de 26.73 milisegundos con respecto a la solución planteada en esta investigación. Esta mínima diferencia entre los tiempos de respuesta se debe principalmente a los procesos de transformación y serialización que realiza el API para presentar los datos en los cuatro diferentes formatos anteriormente mencionados y probados.

Las pruebas revelan un correcto funcionamiento de cada una de las partes que conforman la solución que se plantea en el presente trabajo de investigación, demostrado así que los problemas de interoperabilidad de la UTPL pueden ser atendidos implementando soluciones de este tipo.

4. Conclusiones

En este trabajo se presentó un enfoque novedoso y manejable para generar un ambiente de interoperabilidad semántica en universidades, basado en la construcción de grafos de conocimiento a través del enfoque de Web Semántica y las tecnologías de Linked Data. Este marco se presenta como una alternativa viable a los problemas de interoperabilidad de datos y gobierno de datos en una institución de educación superior.

La evaluación muestra que el marco propuesto permite aprovechar el potencial de Linked Data en términos de posibilidades de desambiguación, integración, clasificación, y transformación de datos relacionales provenientes de fuentes heterogéneas y distribuidas, además de acercar a la UTPPL a un nivel óptimo de interoperabilidad semántica a nivel institucional.

Al igual que en las bases de datos relacionales, los grafos de conocimiento (knowledge graph, KG) se pueden utilizar en el back-end de las aplicaciones, y responder consultas estructurales escritas en SPARQL, además, es posible ejecutar consultas federadas basadas en grafos de conocimiento múltiples y que permitan inferir conocimiento clave en la toma de decisiones. El habilitador de esta característica proviene del hecho de que los datos y su semántica se representan explícitamente usando grafos de conocimiento, en lugar de estar en las capas de la lógica de negocios de las aplicaciones.

Como parte de trabajos futuros, los autores se encuentran explorando diferentes formas para extender el marco descrito e integrar el soporte del razonamiento, de modo que, por un lado, se pueda razonar en tiempo real con los datos y, por otro lado, el propio razonamiento se puede utilizar para inferir consultas automáticamente. Finalmente, los autores recomiendan a los responsables de la gestión de datos de las instituciones de educación superior, la adopción de estándares, protocolos y tecnologías abiertas establecidas por la industria tecnológica, así como la adopción de buenas prácticas en la gestión de datos, con el fin de garantizar un alto nivel de interoperabilidad semántica.

Referencias

1. Ieee recommended practice for architectural description of software-intensive systems. IEEE Std 1471-2000 pp. i–23 (2000)
2. Barata, J., Cunha, P.R.: Synergies between quality management and information systems: a literature review and map for further research. *Total Quality Management & Business Excellence* 28(3-4), 282–295 (2017), <https://doi.org/10.1080/14783363.2015.1080117>
3. Bernres-Lee, T.: Design Issues: Linked Data (2006), <https://www.w3.org/DesignIssues/LinkedData.html><http://www.w3.org/DesignIssues/LinkedData.html>
4. Cadme, E., Piedra, N.: Producing linked open data to describe scientific activity from researchers of Ecuadorian universities. In: 2017 IEEE 37th Central America and Panama Convention (CONCAPAN XXXVII). pp. 1–6 (nov 2017)

5. Chicaiza, J., López, J., Piedra, N., Martínez, O., Tovar, E.: Usage of social and semantic web technologies to design a searching architecture for software requirement artefacts. *Software, IET* 4, 407–417 (2011)
6. Garcia-Gonzalez, H., Gayo, J.E.L., Paule-Ruiz, M.: Enhancing e-Learning content by using Semantic Web technologies. *IEEE Transactions on Learning Technologies* (2016)
7. Ibrahim, A., Bucur, A., Dekker, A., Marshall, M.S., Perez-Rey, D., Alonso-Calvo, R., Stenzhorn, H., Yu, S., Krykwinski, C., Laarif, A., Others: Analysis of the suitability of existing medical ontologies for building a scalable semantic interoperability solution supporting multi-site collaboration in oncology. In: *Bioinformatics and Bioengineering (BIBE), 2014 IEEE International Conference on*. pp. 204–211. IEEE (2014)
8. Khalid, N., Pasha, M., Rehman, S.U., Ahmad, H.F., Suguri, H.: Ontology Services between Agents and OWL Based Web Services. In: *Semantics, Knowledge and Grid, Third International Conference on*. pp. 176–181 (oct 2014)
9. Kríž, V., Hladka, B.: RExtractor: a Robust Information Extractor. In: *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Demonstrations*. pp. 21–25. Association for Computational Linguistics (2015), <http://www.aclweb.org/anthology/N15-3005>
10. Kruchten, P.B.: The 4+1 view model of architecture. *IEEE Software* 12(6), 42–50 (Nov 1995)
11. Piedra, N., Chicaiza, J., López, J., Tovar, E., Martinez, O.: Finding OERs with Social-Semantic search. In: *2011 IEEE Global Engineering Education Conference (EDUCON)*. pp. 1195–1200 (2011)
12. Piedra, N., Chicaiza, J., López, J., Tovar, E.: Seeking Open Educational Resources to Compose Massive Open Online Courses in Engineering Education An Approach based on Linked Open Data. *Journal of Universal Computer Science* 21(5), 679–711 (may 2015)
13. Piedra, N., Chicaiza, J., Lopez-Vargas, J., Caro, E.T.: Guidelines to producing structured interoperable data from Open Access Repositories. In: *Proceedings - Frontiers in Education Conference, FIE*. vol. 2016-Novem, pp. 1–9 (oct 2016)
14. Piedra, N., Chicaiza, J., Quichimbo, P., Saquicela, V.c., Cadme, E., LÁ\textthreesuperiorpez, J., Espinoza, M., Tovar, E.: Marco de Trabajo para la Integración de Recursos Digitales Basado en un Enfoque de Web Semántica. *RISTI - Revista Ibérica de Sistemas e Tecnologías de Información* pp. 55–70 (2015), http://www.scielo.mec.pt/scielo.php?script=sci_arttext&pid=S1646-98952015000100006&nrm=iso
15. Piedra, N., Suárez, J.: SmartLand-LD: A Linked Data approach for Integration of heterogeneous datasets to intelligent management of high biodiversity territories (2018)
16. Piedra, N., Tovar, E., Colomo-Palacios, R., Lopez-Vargas, J., Alexandra Chicaiza, J.: Consuming and producing linked open data: the case of Opencourseware. *Program* 48(1), 16–40 (2014)
17. Sonsilphong, S., Arch-int, N.: Semantic Interoperability for Data Integration Framework using Semantic Web Services and Rule-based Inference: A case study in healthcare domain. *Journal of Convergence Information Technology* 8(3) (2013)
18. Zaveri, A., Rula, A., Maurino, A., Pietrobon, R., Lehmann, J., Auer, S.: Quality Assessment for Linked Open Data: A Survey. *Semantic Web Journal* (by IOS Press) 1, 1–31 (2014), <http://www.semantic-web-journal.net/content/quality-assessment-methodologies-linked-open-data>