




Arquitectura para la Creación y Enriquecimiento Automático de Ontologías a partir de Datos Enlazados

Architecture for the Automatic Creation and Enrichment of Ontologies from Linked Data

a Ricardo José Dos Santos -Guillén, b Eduard Gilberto Puerto -Cuadros
c Jose Lisandro Aguilar.

-  a PhD(c). en Ciencias Aplicadas, Centro de Estudios en Microelectrónica y Sistemas Distribuidos (CEMISID), ricardojds@gmail.com, Universidad de Los Andes, Mérida, Venezuela. ORCID: 0001-8752-8550
-  b PhD. en Ciencias Aplicadas, Facultad de Ingeniería, Departamento de Sistemas e Informática, Grupo GIA, eduardpuerto@ufps.edu.co Universidad Francisco de Paula Santander, Cúcuta-Colombia. ORCID: 0000-0001-9361-5837
-  c PhD. Ciencias de la Computación Centro de Estudios en Microelectrónica y Sistemas Distribuidos (CEMISID), aguilarjos@gmail.com, Universidad de Los Andes, Mérida, Venezuela. ORCID: 0003-4194-6882.

Recibido: Julio 1 de 2021 Aceptado: Noviembre 8 de 2021

Forma de citar: Ricardo José Dos Santos-Guillén, b E.G Puerto-Cuadros J.L Aguilar. "Arquitectura para la Creación y Enriquecimiento Automático de Ontologías a partir de Datos Enlazados.", *Mundo Fesc*, vol. 12 (24), pp. 8-24, 2022

Resumen

En los ambientes inteligentes (p.ej. ciudades inteligentes) se presentan problemas muy diversos que deben ser atendidos inmediatamente, por lo tanto, estos ambientes deben generar conocimiento que les permitan responder a esas necesidades particulares, aprovechando la información del contexto (ej.: datos históricos, medición de sensores, descripción del problema, entre otras). Como propuesta para la generación oportuna de estas bases de conocimiento, en este trabajo se desarrolla una arquitectura que permite crear y enriquecer ontologías emergentes de forma autónoma, usando como insumo el paradigma de Datos Enlazados, que son estructuras de datos que se vinculan unos con otros para servir tanto a usuarios humanos como a otros sistemas dentro de la web semántica. En este trabajo se especifican los servicios que ofrecen la capacidad de generar ontologías emergentes según el problema que se presente, buscando aprender los distintos axiomas (conceptos o propiedades) de las estructuras que posean las fuentes de Datos Enlazados empleada. Luego, estas ontologías creadas son enriquecidas/pobladas con la información que se extraiga desde las distintas fuentes de datos enlazados, para ser devueltas al sistema solicitante de forma automática. Para el desarrollo de este trabajo se usa la metodología MEDAWEDE, que permite especificar esta arquitectura, y guiar la construcción de los servicios para consumir las fuentes descritas como datos enlazados. De dicha metodología se utilizan sus seis etapas: i. Especificación: para analizar y seleccionar las fuentes de datos ontológicas a usar. ii. Modelado: para implementar la base de datos ontológica del área de estudio de las ontologías a generar. iii. Generación: para realizar el proceso de transformación, filtrado e integración de los datos en la ontología generada. iv. Vinculación: para asociar los datos de la ontología generada con el conocimiento proveniente de los conjuntos de datos enlazados. v. Publicación: para poner a disposición la ontología generada en los distintos formatos estándares. vi. Explotación: esta etapa se considera porque está dedicada a definir interfaces para acceder a la ontología. Finalmente, el artículo presenta un caso de estudio enfocado en mostrar el proceso de generación de una ontología emergente en el área de la pandemia del Covid-19, donde se aprovecha la información disponible en la Internet a través de ontologías preexistentes y fuentes de datos enlazados.

Palabras clave: Ontología Emergente, Arquitectura, Datos enlazados, Alineación, Servicio Web.

Autor para correspondencia: *Correo electrónico:
maritzacarolinajm@ufps.edu.co



© 2022. Fundación de Estudios Superiores Comfanorte.

Abstract

In smart environments (e.g.: smart cities) there are very diverse problems that must be addressed immediately; therefore, these environments must generate knowledge that allows them to respond to those particular needs, taking advantage of information from the context (e.g.: historical data, sensor measurement, problem description, among others). As a proposal for the timely generation of these knowledge bases, in this work an architecture is developed for allows to create and enrich emerging ontologies autonomously, using as input the linked data paradigm, which are data structures that are linked to each other. to serve both human users and other systems within the semantic web. This work specifies the services that offer the ability to generate emerging ontologies according to the problem that arises, seeking to learn the different axioms (concepts or properties) of the structures that the linked data sources used have. Then, these created ontologies are enriched/populated with the information that is extracted from the different linked data sources, to be returned to the requesting system automatically. For the development of this work, the MEDAWE-DE methodology is used, which allows specifying this architecture, and conducting the construction of services to consume the sources described as linked data. Its six stages are used from this methodology: i) Specification: to analyze and select the ontological data sources to use. ii) Modeling: to implement the knowledge model of the study area of the ontologies to be generated. iii) Generation: to carry out the process of transformation, filtering, and integration of the data in the generated ontology. iv) Linking: to associate the data from the generated ontology with the knowledge from the linked data sets. v) Publication: to make the generated ontology available in the different standard formats. The exploitation stage is considered because it is dedicated to defining interfaces to access the ontology. The article also presents a case study focused on showing the process of generating an emerging ontology in the area of the Covid-19 pandemic, where the information available on the Internet is used through pre-existing ontologies and linked data sources.

Keywords: Migration, immigration, social work, human rights, border.

Introducción

La integración del conocimiento a partir de fuentes dispersas de datos, es un problema que se tiene para el funcionamiento oportuno y eficaz de Ambientes Inteligentes (AmI) a la hora de tomar decisiones, dado que requieren de conocimiento que le permitan garantizar que la respuesta sea acorde al dominio que se les plantea. Para ello, estos AmI requieren de datos que generalmente son tomados desde sensores dispuestos dentro del ambiente. Sin embargo, dentro de estos AmI se requieren, a su vez, información histórica relevante; es decir, aquella que los sensores han recolectado con antelación o que está albergada en fuentes de datos externas, como son las Ontologías y los Datos Enlazados.

Según [1],[2],[3] una ontología es la representación del conocimiento sobre un dominio específico, donde un conjunto de objetos y relaciones interactúan de forma consistente y coherente. Es decir, un conjunto de axiomas lógicos diseñados para explicar el significado previsto de un vocabulario que pueden ser procesado por máquinas para usarse en algunos tipos de aplicaciones [4],

[5], [6],[7]. Así mismo, los Datos Enlazados, según [8] describen una forma de publicar los datos estructurados para que se puedan interconectar entre ellos. Los autores, a su vez citando a [9] añaden que los Datos Enlazados describen un conjunto de prácticas para publicar, compartir y conectar piezas de datos, información y conocimiento en la Web Semántica, usando identificadores URIs y RDF para describir los recursos publicados.

Por otra parte, dado que estos AmI son un entorno donde coexisten las personas con las máquinas de una manera cómoda e imperceptible y estas se ponen a su servicio para facilitarles la vida y hacerlas más agradable, en donde convergen computadores ubicuos incrustados en objetos cotidianos, comunicaciones inalámbricas entre ellos, interfaces de nueva generación, sensores biométricos, agentes inteligentes, sistemas de personalización, entre otras cosas, se requieren de Servicios Web dispuestos a responder las necesidades que se van generando en estos ambientes.

Los Servicios Web son aplicaciones auto contenidas y auto descriptivas que pueden ser alojadas,

publicadas e invocadas a través de la web [10], [11]. Con esto, los autores definen un servicio web como un objeto de software que puede ser ejecutado a través de internet usando protocolos estandarizados, con el fin de cumplir tareas, funciones o ejecutar procesos completos de negocio. Por otro lado, sirviendo de complemento a la definición anterior, en [12] explican que los servicios web, y en general las arquitecturas orientadas a servicios, son tecnologías que emergieron como elección al querer implementar sistemas distribuidos a través de la modularidad. Gracias a esto, el negocio se divide en pequeñas funciones especializadas que se publican como servicios.

Actualmente, existen diversos trabajos orientados a responder a estos y otros problemas relacionados con los AmI [13]. El primero de estos trabajos [14] propone un nuevo método de enriquecimiento de ontologías a partir de extracción de conceptos. Este proceso semi automático puede exportar datos enlazados minando textos, sin importar su lenguaje, estilo o dominio basado en redes neuronales [15]. Por otro lado, otro trabajo [4] detalla el uso de datos abiertos para impulsar la innovación y el crecimiento económico de ciudades inteligentes; por lo que proyectan una arquitectura que sirve como mediador entre la explotación de datos abiertos y la generación de conocimiento. [16] destaca la importancia que tiene la metadata dentro del mundo de los datos enlazados al momento de publicar datos enlazados de manera completa y consistente. En otro orden de ideas, con una temática similar a la del proyecto actual, los autores de [17] muestran la importancia para manejar los datos enteros de una ciudad en forma de datos enlazados. Proponen una metodología para lograr la homogeneidad de los datos, utilizando técnicas de regresión como estrategia para depurar los datos incluyendo datos enlazados integrados. Así mismo, en [18] proponen crear y poblar ontologías a partir de textos no estructurados. Su metodología consiste en la utilización del contexto para mejorar la comprensión del texto de entrada basándose en la búsqueda y análisis de grandes conjuntos de datos enlazados. Finalmente, en [19] se propone un framework para la integración de ontologías, donde se identifican términos comunes para extraer datos similares de distintas fuentes, para relacionarlos dentro de una misma ontología.

Se plantea así la hipótesis que aprovechando los principios de desarrollo de la web 3.0 es posible hacer que este conocimiento sea de fácil acceso

para los sistemas inteligentes nuevos y actuales. Para lo cual se genera la siguiente interrogante ¿Cómo desarrollar un servicio web que permita crear y enriquecer ontologías de forma autónoma, usando fuentes de datos enlazados para que sirva para la toma de decisiones en AmI? ¿De qué manera generar ontologías de forma automática, a partir de datos enlazados en fuentes públicas?.

A continuación, el artículo en su Sección de Materiales y métodos se describe el diseño de la arquitectura computacional propuesta para la creación y enriquecimiento automático de ontologías a partir de datos enlazados siguiendo la metodología MEDAWEDE. En la sección de Resultados y discusión, se presenta un caso de entorno en el dominio del Covid-19, donde se presenta el uso de la arquitectura. Finalmente se dan las Conclusiones en torno a los resultados.

Materiales y métodos

El diseño arquitectónico de la propuesta se basa en la metodología MEDAWEDE [20], que permite especificar y guiar la construcción de los servicios para consumir las fuentes descritas como datos enlazados. La metodología tiene seis etapas: i. Especificación: para analizar y seleccionar las fuentes de datos ontológicas a usar. ii. Modelado: para implementar la base de datos ontológica del área de estudio de las ontologías a generar. iii. Generación: para realizar el proceso de transformación, filtrado e integración de los datos en la ontología generada. iv. Vinculación: para asociar los datos de la ontología generada con el conocimiento proveniente de los conjuntos de datos enlazados. v. Publicación: para poner a disposición la ontología generada en los distintos formatos estándares. vi. Explotación: esta etapa se considera porque está dedicada a definir interfaces para acceder a la ontología.

La arquitectura de la solución está constituida por tres capas (ver figura 1). La primera capa, denominada Servidor Web, se encarga de recibir las peticiones web que realicen los clientes. La segunda capa, Explorador de Recursos, consulta los diferentes repositorios de Datos Enlazados a partir del dominio solicitado por el cliente. La tercera capa, Generador Ontológico, procesa los datos de la capa anterior y genera una ontología extendida con Datos Enlazados según el conocimiento solicitado.

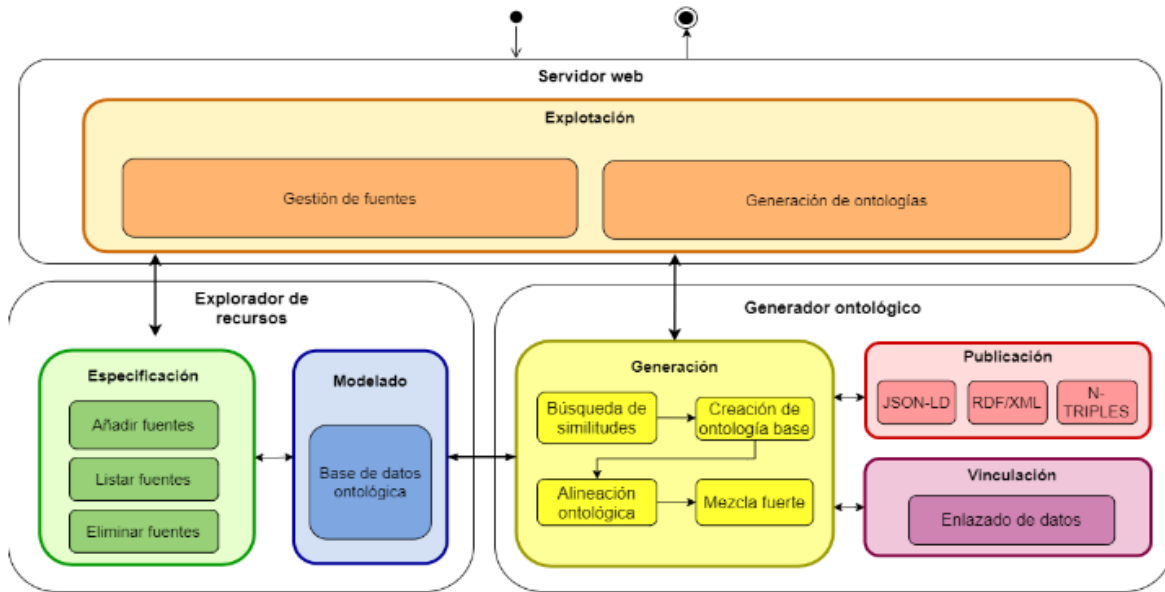


Figura 1 Diagrama de componentes de nuestra arquitectura basada en MEDAWEDE. Fuente Autores.

A continuación, se describe con más detalle cada uno de los componentes de las capas de la arquitectura.

Capa Servidor Web.

Esta capa está compuesta por las interfaces Gestión de Fuentes y Generación de Ontologías. La primera interfaz de Gestión de Fuentes tiene como objetivo ofrecer tres servicios: añadir, eliminar y listar las fuentes de datos. La segunda interfaz tiene por objetivo ofrecer el servicio que genera la ontología enriquecida con Datos Enlazados explotando el conocimiento dispuesto por la primera interfaz.

Capa Explorador de recursos.

La capa del Explorador de recursos está constituida por los componentes de Especificación y modelado. En el componente de Especificación se gestionan las ontologías como fuentes de conocimiento, donde se seleccionan y consultan los conjuntos de datos en los repositorios disponibles ya sea localmente o desde Internet. En el componente de Modelado se genera una Base de Datos Ontológica (BDO) a partir del conjunto de ontologías del componente anterior.

Capa Generador ontológico. Esta capa está conformada por tres componentes: El componente de Generación, de Vinculación y de Publica-

ción (ver Figura 2). En Generación, se consultan conceptos ontológicos a partir del dominio solicitado, para poblar una ontología nueva que será el resultado de la alineación y mezcla del conjunto de ontologías de la BDO que correspondan a la consulta.

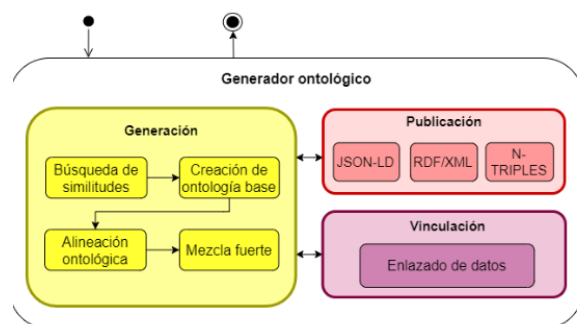


Figura 2. Componentes Capa Generador Ontológico. Fuente Autores.

A continuación, se describe con más detalle el proceso: Un primer paso consiste en la búsqueda de similitudes, donde se reciben los términos de búsqueda que se usan como criterio para la consulta a la BDO y se seleccionan los conceptos ontológicos que contienen estos términos. El siguiente paso es la creación de la ontología base, que corresponde a una ontología nueva que contiene únicamente las clases de la búsqueda.

Esta ontología albergará también los conceptos que sean seleccionados en los siguientes pasos. A continuación, sigue el proceso de alineación ontológica. Durante este paso se realiza el preprocesado de los conceptos ontológicos obtenidos en la búsqueda, buscando ponderar su relación con cada uno de los términos de búsqueda. Estas ponderaciones ayudan al filtrado de los conceptos que no corresponden a la alineación ontológica, buscando asegurar la calidad de la ontología generada. El paso final durante el proceso de generación es la mezcla fuerte. Durante este se copian los conceptos seleccionados de las ontologías de la BDO y se integran dentro de la ontología base, obteniendo como resultado una ontología generada y poblada que corresponde al conocimiento del dominio solicitado en la búsqueda.

En Vinculación se enriquece la ontología generada en el componente anterior con Datos Enlazados. Para ello, se usa el servicio de búsqueda de DBpedia (<https://lookup.dbpedia.org/api/search/>) para enriquecer con información de DBpedia a cada concepto de la ontología. En Publicación se transforma la ontología enriquecida con Datos Enlazados al formato requerido por la interfaz. Entre los formatos que actualmente maneja son JSON-LD, RDF/XML y N-TRIPLES.

A continuación, se desarrolla un caso de estudio a partir del cual se evalúa la arquitectura.

Resultados y discusión

Para el desarrollo y prueba de la arquitectura propuesta (ver figura 1) se ha tomado como fuente o insumo de conocimiento el marco ontológico de sobre el dominio del Covid-19. Para esto se ha seleccionado un conjunto de ontologías vinculadas al Covid-19 del repositorio de ontologías <https://bioportal.Bioontology.Org/ontologies> (ver Tabla 1).

N	Ontología	Descripción	URL
1	COviD-19 Ontology for Cases and Patient information (CODO)	Contiene conocimiento relacionado con el paciente y el Covid-19 especificado en 52 conceptos	https://bioportal.bioontology.org/ontologies/CODO
2	COVID-19 Surveillance Ontology (COVID19)	Contiene conocimiento relacionado con el control, vigilancia del Covid-19 especificado en 52 conceptos.	https://bioportal.bioontology.org/ontologies/COVID19
3	The COVID-19 Infectious Disease Ontology (IDO-COVID-19)	Contiene conocimiento relacionado con enfermedades infecciosas entorno al Covid-19 especificado en 486 Conceptos	https://bioportal.bioontology.org/ontologies/IDO-COVID-19
4	COVID-19 Ontology (COVID-19)	Contiene conocimiento general del Covid especificado en 2286 conceptos	https://bioportal.bioontology.org/ontologies/COVID-19
5	WHO COVID-19 Rapid Version CRF semantic data model (COVIDCRFRAPID)	Contiene conocimiento sobre datos clínicos estandarizados sobre Covid-19 especificado en 398	https://bioportal.bioontology.org/ontologies/COVID-CRFRAPID

Tabla 1. Fuentes de Conocimiento Ontológicas sobre Covid-19 Fuente: Autores

Con estas fuentes de datos se establece la especificación y el modelado de la base de datos ontológica del Servicio Web. A continuación, se ha hecho la solicitud al servicio web que genere una ontología en formato RDF/XML acerca de los términos de búsqueda “Test” y “Covid”.

Para llevar a cabo el paso de Generación, la aplicación ha debido buscar en diversas fuentes sinónimos de cada término de búsqueda. Por ejemplo, para “Test” ha encontrado sinónimos como “trial, exam, quiz” entre otros, mientras que para “Covid” ha encontrado “Coronavirus”. Cabe aclarar en este punto que el caso de prueba se realizó en inglés por ser el idioma de las ontologías utilizadas. El sistema podrá ser expandido en un futuro para funcionar de la misma manera con otros idiomas. Continuando con el proceso, se utiliza la librería de Python OlwReady 2 para realizar consultas

sobre la base de datos ontológica, donde se obtienen conceptos ontológicos de cualquiera de las ontologías integradas, donde sus etiquetas contengan cualquiera de los términos de búsqueda o sus sinónimos. Entre estos resultados destacan algunos como “Untested for Covid-19”, “Covid-19 Diagnosis”, “Tested for 2019-nCov (Wuhan) infection”.

Sin embargo no todos estos conceptos ontológicos terminarán poblando la ontología, puesto que muchos de ellos no guardan relación semántica con el dominio que se consultó, como es el caso de resultados como “intestine cancer”, “assay screened entity”, “testis”; por lo tanto es necesario realizar un filtrado entre los conceptos. Su objetivo es ponderar cada concepto ontológico obtenido en la consulta de acuerdo a su relación con los términos de búsqueda. Para esto se preprocesan y comparan sus etiquetas simplificadas, otorgando mayor puntaje si el concepto coincide con varios términos de la búsqueda a la vez.

Contextualizando este paso al caso de prueba, un concepto cuya etiqueta sea “Tested for covid-19” tendrá una puntuación mayor a uno cuya etiqueta sea “Study for covid propagation”, ya que en el primer caso se hace referencia a “Test” y a “Covid” a la vez, mientras que en el segundo sólo se hace referencia a “Covid”. De la misma forma, un concepto etiquetado como “intestine cancer” obtendrá una puntuación muy baja ya que sólo se relaciona con “Test” como una subcadena dentro de uno de sus términos (“intestine”)

Esto se aprecia mejor en la figura 3 donde se ilustran algunas comparaciones entre diversos tokens provenientes del preprocesado de las etiquetas de conceptos ontológicos distintos (Azul) con los conceptos de búsqueda y sus sinónimos (Verde oscuro) El primer caso, la comparación de “covid” como token en el concepto evaluado con “Covid” como término de búsqueda es una coincidencia perfecta, por tanto su puntuación será máxima (1). En el segundo caso se observa que el término “Test” está inmerso parcialmente dentro del token “untest” con un sufijo o prefijo añadido, por esto su puntuación es la mitad de la de las coincidencias perfectas (0.5). En el caso de que el término de búsqueda esté inmerso completamente como una subcadena del token del concepto, siendo el caso de “Test” e “intestin”, la ponderación es de un cuarto de la coincidencia perfecta (0.25), buscando castigar posibles conceptos sin relación semántica con la búsqueda. En caso de los sinónimos de los términos de búsqueda (verde claro) se utiliza el mismo criterio, pero la puntuación obtenida se reducirá a la mitad.

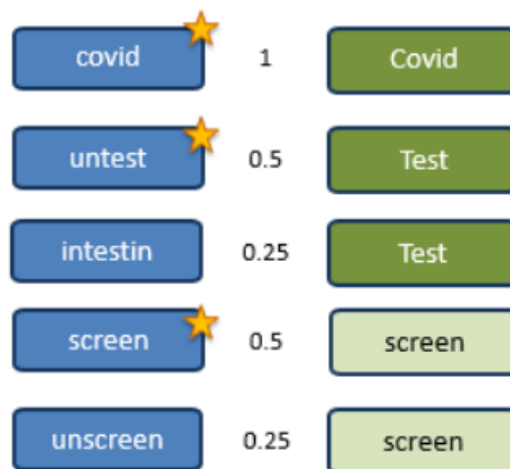


Figura 3. Criterios de ponderación al comparar conceptos ontológicos dentro de los filtros de búsqueda. Fuente Autores.

Después de la ponderación se escogen los conceptos que cumplan con rangos de puntajes, teniendo en cuenta el umbral de aceptación, que es un porcentaje escogido por quien solicita la ontología, que define qué tan severo o permisivo es el filtro de los conceptos. Finalmente, habiendo concluido la alineación ontológica, sigue el subproceso de mezcla fuerte, donde todos aquellos conceptos que superen los filtros se integran dentro de la ontología base recién creada.

Concluido esto, puede procederse al paso de Vinculación de la metodología, donde se buscan conceptos en los Datos Enlazados que puedan ser equivalentes a los de la ontología generada, permitiendo vincularse de forma externa, mientras que el razonador interno de manera automática busca equivalencias y relaciones entre los conceptos de la propia ontología. En el caso del ejemplo, la búsqueda de recursos en dbpedia respecto a “Covid” ha permitido la vinculación entre el concepto covid de la ontología generada (covid_test#covid) con el concepto Coronavirus_disease_2019 de dbpedia (http://dbpedia.org/resource/Coronavirus_disease_2019)

Una vez la ontología se ha generado y vinculado, puede publicarse en el formato seleccionado, en este caso RDF/XML. La figura 4 muestra una representación gráfica (<http://www.visualdataweb.de/webvowl/>) de la ontología generada resultante del caso de prueba, haciendo énfasis en los conceptos raíz de la búsqueda (“covid”, “test”, “covid_test”) y su relación con los conceptos obtenidos en la búsqueda.

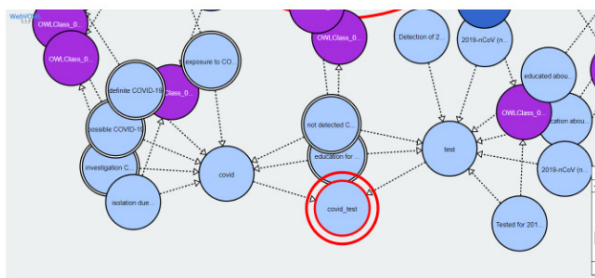


Figura 4. Fragmento de los Conceptos filtrados y vinculados en la ontología generada en el caso de prueba. Fuente Autores.

Conclusiones

El artículo presenta un servicio web enriquecedor de ontologías que se encargue de poblar ontologías con nuevos conceptos y relaciones, a partir de Datos Enlazados de forma automática. Un servicio que facilita la integración del conocimiento a partir de fuentes dispersas de datos enlazados, aprovechando los principios de desarrollo de la web 3.0 de fácil acceso para los sistemas inteligentes nuevos y actuales que lo soliciten. Se ofrece una solución la cual es desplegada en forma de servicio web, para ser utilizado de forma remota.

El algoritmo propuesto para la creación de ontologías permite que la respuesta sea acorde al dominio que se le plantea, a través de filtros en la extracción e integración de datos desde repositorios dispersos y otras ontologías existentes; todo con el fin de que el conocimiento generado sea aprovechado al máximo por los sistemas siguientes dentro del ciclo de vida de los datos.

Se ofrece una solución capaz de servir, de forma general, a arquitecturas no supervisadas, de sistemas inteligentes que puedan consumir este servicio web sin necesidad de participación humana. De ahí que la arquitectura propuesta tenga propiedades de autonomía, para garantizar una respuesta óptima, que no se obtendría al existir intervención humana en el proceso de selección de datos o de integración del conocimiento.

Se realizó un caso de prueba de enriquecimiento en el dominio de COVID para consumir este servicio es el middleware MiSCi, propuesto por Dos Santos (2020).

Cabe notar que el sistema desarrollado no servirá como repositorio para albergar ontologías y/o conjuntos de datos enlazados. El sistema será alojado en un servidor de pruebas y su despliegue final no contempla alojamiento en producción ni nombre de dominio. El proyecto utiliza como datos fuentes de Datos Enlazados.

Siguiendo metódicamente la arquitectura diseñada, se especificó un caso de estudio enfocado al dominio del Covid-19, donde se obtuvo de manera automática una ontología enriquecida con Datos Enlazados para este dominio. La arquitectura mostró que se pueden obtener ontologías no supervisadas lo cual permite resolver situaciones emergentes de manera automática en los AmI, permitiendo la interoperabilidad y razonamiento entre los diferentes componentes del AmI.

Es de notar que en la alineación se obtuvo algunas clases disjuntas que no tienen relaciones

fuertes con los demás conceptos, pero aun así obtuvieron la puntuación suficiente para superar los filtros o umbrales. En la revisión del enlazado de conceptos automático con DBpedia, por su parte, se presentaron inconsistencias semánticas. Estas debilidades serán abordadas en trabajos futuros.

Referencias.

- [1] T. R. Gruber, N. Guarino y R. Poli, "Formal ontology in conceptual analysis and knowledge representation," capítulo "Towards principles for the design of ontologies used for knowledge sharing" en *Conceptual Analysis and Knowledge Representation*, 1993.
- [2] I. Osman, S. B. Yahia, and G. Diallo, "Ontology integration: approaches and challenging issues," *Information Fusion*, vol. 71, pp. 38-63, 2021.
- [3] A. González-Eras, R. Dos Santos, J. Aguilar, and A. Lopez, "Ontological engineering for the definition of a COVID-19 pandemic ontology," *Informatics in Medicine Unlocked*, vol. 28, 100816, 2022.
- [4] O. El Bacha, O. Jmad, Y. El Bouzekri El Idrissi y N. Hmina, "Exploiting Open Data to Improve the Business Intelligence & Business Discovery Experience". no 27 In *Proceedings of the 2nd international Conference on Big Data, Cloud and Applications*, (New York), pp 1–6, March 2017.
- [5] A. Degbelo, "A snapshot of ontology evaluation criteria and strategies," in *Proceedings of the 13th International Conference on Semantic Systems*, septiembre de 2017, pp. 1-8.
- [6] N. Guarino (Ed.), *Formal ontology in information systems: Proceedings of the first international conference (FOIS'98)*, June 6-8, Trento, Italy, vol. 46, IOS press, 1998.
- [7] N. Piedra, J. Chicaiza, E. Cadme y R. Guaya, "Una aproximación basada en Linked Data para la detección de potenciales redes de colaboración científica a partir de la anotación semántica de producción científica: Piloto aplicado con producción científica de investigadores ecuatorianos," *Maskana*, vol. 5, 2014.
- [8] T. Rodríguez, R. D. Santos y J. Aguilar, "Metodología para el desarrollo de aplicaciones Web

utilizando datos enlazados,” in Conferencia Nacional de Computación, Informática Y Sistemas (CoNCISa 2017), vol. 5, pp. 978-980, 2017.

[9] O. Hartig y A. Langegger, “A database perspective on consuming linked data on the web,” *Datenbank-Spektrum*, vol. 10, pp. 57-66, octubre de 2010.

[10] D. Fensel y C. Bussler, “The web service modeling framework WSMF,” *Electronic Commerce Research and Applications*, vol. 1, no. 2, pp. 113-137, 2002.

[11] F. H. Vera-Rivera, E. Puerto, H. Astudillo y C. M. Gaona, “Microservices backlog—A genetic programming technique for identification and evaluation of microservices from user stories,” *IEEE Access*, vol. 9, pp. 117178-117203, 2021.

[12] B. Benatallah, F. Casati, and F. Toumani, “Analysis and Management of Web Service Protocols,” in P. Atzeni, W. Chu, H. Lu, S. Zhou, and T. W. Ling (Eds.), *Conceptual Modeling – ER 2004*, vol. 3288, *Lecture Notes in Computer Science*, Springer, Berlin, Heidelberg.

[13] M. Sánchez, J. Aguilar y E. Exposito, “Integración SOA-MAS en ambientes inteligentes,” *Dyna*, vol. 85, no. 206, pp. 268-282, 2018.

[14] A. Alba, A. Coden, A. L. Gentile, D. Gruhl, P. Ristoski y S. Welch, “Multi-lingual concept extraction with linked data and human-in-the-loop”, no 24 In *Proceedings of the Knowledge Capture Conference*, pp. 1-8, (New York), December 2017.

[15] J. M. V. Carrero y E. Puerto, “GeoMotor: Design with nature. Recognition of geometries using a convolutional neural-network approach (CNN),” in *ICGG 2020-Proceedings of the 19th International Conference on Geometry and Graphics*, Springer International Publishing, 2021, pp. 916-919.

[16] A. Dimou, T. De Nies, R. Verborgh, E. Mannens, P. Mechant y R. Van de Walle, “Automated metadata generation for Linked Data generation and publishing workflows,” in *Proceedings of LDOW2016*, CEUR-WS.org, 2016, pp. 1-10.

[17] S. Bischof, C. Martin, A. Polleres y P. Schneider, “Collecting, integrating, enriching and republishing open city data as linked data,” in *The Semantic Web-ISWC 2015: 14th International Semantic Web Conference*, Bethlehem, PA, USA, October 11-15, 2015, *Proceedings, Part II*, vol. 14, pp. 57-75, Springer International Publishing.

[18] M. Booshehri y P. Luksch, “Towards adding linked data to ontology learning layers,” in *Proceedings of the 16th International Conference on Information Integration and Web-based Applications & Services*, diciembre de 2014, pp. 401-409.

[19] L. Zhao y R. Ichise, “Ontology integration for linked data,” *Journal on Data Semantics*, vol. 3, pp. 237-254, 2014.

[20] R. Dos Santos, J. Aguilar, and E. Puerto, “A Meta-Learning Architecture based on Linked Data,” in *Proc. XLVII Latin American Computing Conference (CLEI)*, 2021.