

Un método de evaluación de ontologías basado en el análisis de categorías de términos.

An Ontology evaluation method based on terms category analysis.

Presentación: 17/10/2023

Christian German Schwab

Universidad Regional Santa Fe, Facultad Regional Santa Fe.
schwab1024@yahoo.com.ar

Resumen

En los proyectos de implementación de ontologías es habitual la consideración de ontologías externas para su incorporación en el mismo. La forma en que la importación de ontologías ocurre en un proyecto es por medio de su inclusión, modificación, o extensión a efecto que responda lo mejor posible al objetivo y requerimientos que la ontología en desarrollo tiene propuestos. En este trabajo se expone un método para evaluar una ontología respecto de un conjunto de categorías de términos con tal de poder responder si una ontología es factible de ser incorporada al proyecto. El método basa sus conclusiones en la cantidad de términos que se puedan reconocer en las categorías derivadas de los términos incluidos en las taxonomías definidas en una ontología. Se propone un algoritmo para definir el método y la aplicación en una ontología concreta para ejemplificar un caso de estudio.

Palabras clave: Proyecto, Ontología, Evaluación, Extensión, Categorías, Taxonomías.

Abstract

In ontology implementation projects, it is common to consider external ontologies to add into the project. The way in which the import of ontologies happens in a project is through their inclusion, modification, or extension so that it responds as best as possible to the objective and requirements that the ontology under development has proposed. In this work, a method is presented to evaluate an ontology over a set of categories of terms in order to be able to answer if an ontology is feasible to be incorporated into the project. The method bases its conclusions on the number of terms that can be recognized in the categories derived from the terms included in the taxonomies defined in an ontology. An algorithm is proposed to define the method and an application in a specific ontology to exemplify a test case.

Keywords: Project, Ontology, Evaluation, Extension, Categories, Taxonomies.

Introducción

Nataly Noy define como objetivos principales que justifican el desarrollo de una ontología a las siguientes premisas: “... *compartir un entendimiento común de la estructura de información entre personas o agentes de sistemas ...*”, “*habilitar el reuso del dominio de conocimiento*” y “*analizar el conocimiento del dominio*” (Noy, 2001). De las premisas anteriores se desprende que los objetivos principales del desarrollo de una ontología es analizar el conocimiento del dominio, separar el conocimiento del dominio del conocimiento operativo, en tal forma que pueda ser compartido tanto por un grupo de personas como de un grupo de procesos.

Para abordar los objetivos anteriores es posible realizar uno de dos escenarios. El primero es a través el desarrollo de una ontología desde cero. El segundo es a partir del reúso de ontologías existentes ya sea, mediante la inclusión de otra ontología, o mediante la modificación o extensión de su implementación para producir una nueva ontología. Respecto de esto último, es necesario realizar una evaluación de las ontologías candidatas con tal de terminar cual de todas es la más apta a efectos del objetivo planteado.

En el marco de este trabajo se utilizará el termino evaluación como la tarea de determinar si una ontología es aplicable a un dominio, subdominio o área de conocimiento específico para diferenciarla de la tarea de determinar: si la calidad de la ontología está acorde a sus objetivos, si se usaron buenas prácticas durante su desarrollo o si no tiene defectos en su construcción.

Respecto de la tarea de evaluación de ontologías en los términos anteriores existen ejemplos entre los que se puede citar al trabajo de Zhe y Dong (2013), en contraposición al trabajo de Gomez Perez (2010) y el de López Rodríguez et al. (2018) en los que la evaluación de las ontologías es respecto de su estructura interna o de su calidad.

Zhe y Dong (2013) proponen un método para facilitar la adecuación de extensiones o modificaciones de distintas versiones de la misma ontología. En el trabajo propuesto por Zhe y Dong, se atraviesa el espacio de nombres de clases, atributos y relaciones, con tal de realizar una fusión ordenada entre ambas ontologías. Los autores proponen un índice para medir el nivel de similaridad entre una clase y otra, en términos de la definición de los atributos de ambas clases y este es usado para decidir qué hacer con la clase analizada. Su objetivo está orientado a reducir el esfuerzo de reescritura de las definiciones de las clases implementadas por la ontología en contraposición de lo que se pretende en este trabajo que es poder evaluar a priori si una ontología es relevante para un área de conocimiento específico.

Gomez Perez (2010) propone que la evaluación técnica de una ontología se realice sobre tres aspectos. El primer aspecto es sobre las definiciones de la ontología, el segundo aspecto es sobre el ambiente de desarrollo usado para su implementación y el tercer aspecto es sobre la documentación que describe la ontología. La autora propone que la evaluación se realice chequeando la estructura de la ontología, la sintaxis de las definiciones y finalmente el contenido de las definiciones. Mediante estas tareas el objetivo planteado es obtener conclusiones respecto de la Consistencia, Completitud y Concisión de la ontología.

López Rodríguez et al. (2018) propone un método de evaluación de ontologías basado en el uso de razonadores, una lista de chequeo de los errores comunes encontrados en el diseño de las ontologías y un conjunto de casos de prueba. Mediante la lista de chequeo se evalúa el cumplimiento de los objetivos de la ontología, mediante los casos de prueba se evalúa el cumplimiento de los requerimientos de la ontología y mediante los razonadores se evalúa el modelo ontológico para la detección de errores conceptuales respecto del dominio de aplicación.

Considerando los objetivos citados anteriormente respecto de disponer una herramienta que permita “*habilitar el reúso del conocimiento del dominio*” y “*analizar el conocimiento del dominio*”, se propone un método que permita obtener todo el conocimiento del dominio modelado por una ontología de forma tal que permita evaluar si su implementación es factible en respecto de los objetivos planteados originalmente por el autor de la ontología.

Metodología

En esta sección se presenta una descripción formal del método de evaluación propuesto con tal que quede expuestos todas las partes del mismo.

El método de evaluación involucra el descubrimiento de un conjunto de categorías que son relevantes al dominio que parten del análisis de la ontología. Posteriormente se puede tomar este conjunto como resultado o compararlo con un conjunto de categorías testigo.

Sea una ontología O , el conjunto C con todos los nombres de clases, atributos de O , c_i cada elemento de C , el conjunto S con todas las categorías derivables de O y sean la categoría s_j un elemento perteneciente a S .

Por cada elemento c_i de C realizar:

- Si el significado de c_i es congruente con la definición de alguna categoría s_j al que pueda asociarse entonces debe agregarse el c_i a la categoría s_j .
- Si el significado del c_i no es congruente con ninguna categoría del conjunto S entonces evaluar el significado de c_i con tal de poder crear una nueva categoría s_j , agregar c_i a s_j y luego agregar s_j a S .
- Si no quedan más elementos c_i en C que analizar, entonces terminar.

Cabe señalar que la pertenencia de un elemento a una categoría no es exclusiva por lo que este puede aparecer en varias categorías simultáneamente. Adicionalmente cada categoría nueva debe ser definida en orden de dar congruencia a la recategorización de los términos.

Para una ontología O , sea la cantidad $t_j(s_j, O)$ la cantidad de elementos referidos a una categoría s_j de una ontología O y sea la cantidad $t(O)$ la cantidad de elementos totales de O , entonces la importancia relativa $r_j(s_j, O)$ de la categoría s_j respecto de la ontología O se define como:

$$r_j(s_j, O) = \frac{t_j(s_j, O)}{t(O)}$$

Sea el conjunto de categorías S compuesta por las categorías s_0, s_1, \dots, s_n de una ontología O de tal forma que

$$S(O) = \{ s_0(O), s_1(O), \dots, s_n(O) \}$$

Para una ontología dada O y un conjunto de categorías S entonces se puede obtener un conjunto R de indicadores tales que:

$$R(S, O) = \{ r(s_0, O), r_1(s_1, O), \dots, r_n(s_n, O) \}, \text{ con } s_j \in S(O) \text{ e } j = 1 \dots n$$

Como ejemplo de uso, el método de evaluación de ontologías se aplica a la Ontología de Mercado Eléctrico o de sus siglas en inglés EMO, Electric Market Ontology (Santos et al., 2018). EMO es una ontología de dominio para lograr la interoperabilidad entre diferentes plataformas de simulación de mercados eléctricos. Su objetivo es servir de base para su extensión o modificación en otras ontologías para mercados eléctricos específicos. Para el desarrollo de este caso de estudio se eligió EMO debido a que es una ontología de dominio con relativamente poca cantidad de términos, clases, atributos y relaciones, lo que la convierten en una candidata ideal para ilustrar el análisis de una ontología. En la Tabla 1 se detalla el conjunto de términos de la ontología EMO.

Para conformar un nuevo conjunto de categorías a partir de los términos de la Tabla 1 por cada termino se contrasta su definición contra la definición extraída por un diccionario diferente. Si ambas definiciones son semejantes entre si entonces no hay necesidad de crear una nueva categoría y se toma el nombre de la clase que subsume el término. En caso contrario se puede crear un nuevo nombre de categoría que represente el termino para obtener una visión diferente de la expresada por la ontología. A diferencia de la tarea de la construcción de taxonomías de las primeras etapas del desarrollo de ontologías, este algoritmo propone que los términos tengan instancias en distintas categorías.

Aplicando el algoritmo presentado anteriormente con cada término de la Tabla 1 se puede recategorizar los términos de EMO en las siguientes categorías:

- **Facturación:** Categoría para la descripción de las funciones, procedimientos, actividades y términos relacionadas con la ejecución de contratos del mercado.
En esta categoría se incluyen los términos Precio Máximo, Precio Mínimo y Precio. La elección de la categorización en esta categoría surge porque en las transacciones comerciales de un mercado eléctrico es el precio lo que se toma como valor para contabilizar la facturación realizada en un periodo de tiempo.
- **Mercado:** Categoría para la descripción de los términos involucrados la interacción entre partes para la compraventa de recursos de interés.
En esta categoría se incluyen a los términos que describen los participantes de un mercado tales como Representante, Operador de Mercado, Operador, Participante, así como también los documentos que se firman entre ellos para fijar

las condiciones de las transacciones comerciales de compra/venta tales como Contrato Bilateral, Oferta, Sesión y Tipo de Transacción. También son incluidos en esta categoría el ámbito en la que los participantes y las condiciones se establecen entre estos tales como Mercado y Tipo de Mercado.

Termino	Definición
Precio Máximo	Precio Máximo aplicable a un Área Geográfica
Precio Mínimo	Precio Mínimo aplicable a una Área Geográfica.
Precio	Precio genérico aplicable en cualquier caso.
Representante.	Participante genérico de un mercado eléctrico genérico. Agrupa un conjunto de participantes de escala menor de forma que pueda participar del mercado.
Contrato Bilateral	Representa un contrato bilateral abstracto del mercado eléctrico.
Mercado	Representa un mercado eléctrico genérico. El mercado ibérico español será instancias de esta clase.
Operador de Mercado	Representa un tipo de operador de mercado genérico.
Tipo de Mercado	Representa un tipo de mercado eléctrico genérico.
Oferta	Representa una oferta genérica en el mercado compuesta por un valor de potencia y un precio.
Operador	Representa un operador de mercado eléctrico genérico.
Participante	Representa un participante genérico del mercado eléctrico.
Sesion	Representa un evento de mercado eléctrico genérico.
Tipo de Transacción	Representa un tipo de transacción del mercado eléctrico. Son instancias de esta clase la acción de compra o venta.
Operador del Sistema	Representa un tipo de operador del sistema genérico dentro de un mercado eléctrico genérico.
Área	Área geográfica donde los eventos de mercado eléctrico ocurren.
Periodo	Momento histórico donde los eventos del mercado eléctrico ocurren.

Tabla 1 – Listado de términos de la ontología EMO.

- **Sistema Administrativo:** Categoría para la descripción de las formas de organización humana para la obtención de objetivos de diversas índoles.
 En esta categoría se incluye los términos que incluyen las formas de organización legal que los representantes conforman para la realización de las transacciones comerciales en un mercado eléctrico genérico. Entre las que se enumeran en esta categoría están los términos de Representante, Operador de Mercado, Operador, Participante, Sesión y Operador de Sistema.
- **Topología:** Categoría para la descripción del posicionamiento de las cosas respecto de sus ubicaciones.
 En esta categoría se incluye el concepto de área para describir el ámbito donde se realizan las transacciones comerciales de un mercado eléctrico genérico. Entre los términos que se pueden incluir en esta categoría se encuentra Área.
- **Unidad de medida:** Categoría para la descripción de unidades de medida incluyendo sus submúltiplos y múltiplos.
 En esta categoría se puede incluir el concepto Periodo porque es el único termino que es una unidad de medida genérica.

Todas las categorías anteriores fueron derivadas a partir de tanto: las definiciones de cada termino especificados en la ontología EMO y a partir de las definiciones de diccionarios externos, tales como la RAE.

A continuación se presentan los resultados en la Tabla 2. En la primera columna de la tabla se incluyen los términos de la ontología analizada y al final de la misma, el total de elementos de la ontología, $t(O)$. En la tabla se puede visualizar la distribución de términos respecto de las categorías detectadas. En esta tabla, la presencia de un término dentro de una categoría se muestra como un 0 (no presente) o como un 1 (presente). En la anteúltima fila de cada columna, contiene la

cantidad de términos $t_f(s_j, O)$ que corresponde a la categoría s_j . En la última fila de cada columna se tiene la importancia relativa $r_f(s_j, O)$ de la categoría s_j respecto de la ontología O .

Términos/Categoría	Facturación	Mercado	Sistema Administrativo	Topología	Unidad Medida	
Precio Máximo	1	1	0	0	0	
Precio Mínimo	1	1	0	0	0	
Precio	1	1	0	0	0	
Representante	0	1	1	0	0	
Contrato Bilateral	0	1	0	0	0	
Mercado	0	1	0	0	0	
Operador de Mercado	0	1	1	0	0	
Tipo de Mercado	0	1	0	0	0	
Oferta	0	1	0	0	0	
Operador	0	1	1	0	0	
Participante	0	1	1	0	0	
Sesión	0	1	1	0	0	
Tipo de Transacción	0	1	0	0	0	
Operador del sistema	0	0	1	0	0	
Área	0	0	0	1	0	
Periodo	0	0	0	0	1	
Total	16	3	13	6	1	
		19%	81%	38%	6%	6%

Tabla 2 – Distribución de términos contra categorías de la ontología EMO.

Resultados y discusión

Sobre el caso de aplicación descrito con anterioridad y a partir de la aplicación del método de evaluación en la ontología EMO, se puede decir que la categoría Facturación tiene una relevancia del 19%, la categoría Mercado tiene una relevancia del 81%, Topología tiene una relevancia del 6% y Unidad de Medida tiene una relevancia del 6%. Se dice que la categoría Mercado es más relevante que las otras categorías porque se pudo ubicar mayor cantidad de términos que en las otras categorías. La suposición que se hace de lo anterior es que existió un esfuerzo más grande de describir la categoría de términos incluidos en Mercado que por ejemplo Topología que solo es descrita por un solo término.

Se asume que la relevancia de la categoría dentro de la ontología radica exclusivamente en la cantidad de términos que pueden ser incluidos en ella. Por lo que, para este método, la relevancia individual de cada término es la misma.

Considerando que uno de los objetivos de la ontología estudiada es servir de base para su extensión o modificación en ontologías de mercados eléctricos específicos, también puede decirse que dadas las categorías derivadas de la lista de términos de EMO también puede incorporarse en ontologías que requieran términos relacionados para Mercados genéricos. En mayor medida puede incorporarse en ontologías que requieran describir procesos de Facturación genéricos. En menor medida puede ser incorporada en ontologías que requieran representar términos de Unidad de Medida y términos de Topología.

La ontología EMO puede ser aplicada en proyectos con objetivos distintos al planteado por el analista de la ontología porque los términos implementados en su taxonomía son de carácter amplio y exceden al dominio del Mercado Eléctrico permitiendo que pueda aplicarse a cualquier mercado genérico.

Como trabajos futuros se deja planteado dos posibles modificaciones para la construcción del índice de relevancia para cada categoría. Una posible extensión de este método es aplicar un promedio ponderado en función de la importancia que le da el

analista a cada termino. Otra posible extensión es usar una tabla para proporcionar un índice de importancia por termino con tal que el índice de relevancia de termino dentro de cada categoría fuera más preciso.

Conclusiones

En este trabajo se planteó como objetivo presentar un método para evaluar una ontología que permita determinar la factibilidad de su reuso en otros proyectos de implementación de ontologías. En la práctica lo anterior se expresa en la pregunta si la ontología analizada puede adaptarse a una situación concreta caracterizada por una categoría de términos.

Esto puede ser respondido por el método de evaluación propuesto, en el sentido que en el caso que una ontología se pueda derivar una categoría específica clave en la realización de los objetivos de otra ontología, entonces es factible de incorporación. Lo anterior no obsta que puedan realizarse evaluaciones de orden cualitativo adicionales que permitan ratificar o rectificar según el resultado obtenido.

Puede concluirse adicionalmente que ningún método de evaluación de ontologías es definitivo para todas las situaciones planteadas en el proceso de desarrollo de ontologías y quedan mucho trabajo en este tema para desarrollar.

Referencias bibliográficas

Noy, N.; McGuinness, D (2001). *Ontology Development 101: A Guide to Creating Your First Ontology*. Knowledge Systems Laboratory, http://www.ksl.stanford.edu/KSL_Abstracts/KSL-01-05.html.

Santos, G., Pinto, T., Praça, I., Vale, Z. (2018). "Iberian electricity market ontology to enable smart grid market simulation". *Energy Inform*, 1, article 13. <https://doi.org/10.1186/s42162-018-0018-2>.

Zhe, Cao; Dong, Lui; (2013). "Research on CIM Update and Extension Based on Ontology". *Research Journal of Applied Sciences, Engineering and Technology*, 14, 2561-2567, <http://dx.doi.org/10.19026/rjaset.6.3738>.

Gomez Perez, A.; (1995). "Some Ideas and Examples to Evaluate Ontologies". Proceedings the 11th Conference on Artificial Intelligence for Applications, <http://dx.doi.org/10.1109/CAIA.1995.378808>.

López Rodríguez, Y. A.; Hidalgo-Delgado, Y.; Martínez. N. (2018). "Un método practico para la evaluación de ontologías". Proceedings of the 3rd International Workshop on Semantic Web 2018, <https://ceur-ws.org/Vol-2096/paper7.pdf>.