

Evaluación empírica del mapeo de reglas de negocio para el desarrollo de ontologías

Emiliano Reynares ¹, María Laura Caliusco ¹, María Rosa Galli ^{1,2}

ereynares@frsf.utn.edu.ar, mcaliusc@frsf.utn.edu.ar, mrgalli@santafe-conicet.gov.ar

¹ CIDISI Research Center, UTN Santa Fe, B. Lavaysse 610, S3004EWB Santa Fe, Argentina

² INGAR-UTN-CONICET, Avellaneda 3657, S3002GJC Santa Fe, Argentina

DOI: 10.17013/risti.14.83-99

Resumen: Las metodologías de desarrollo de ontologías usadas como un artefacto de software aún constituyen un campo abierto de investigación. Esto se debe a la inmadurez de las metodologías actuales; donde las más reconocidas se basan en UML por ser éste un lenguaje maduro para los ingenieros de software. Trabajos recientes proponen el mapeo de expresiones SBVR a sentencias ontológicas como técnica de desarrollo, pero hasta el momento existe sólo un estudio empírico que compara técnicas basadas en SBVR y UML. El presente trabajo analiza la factibilidad técnica del mapeo SBVR a OWL 2 presentando una primera réplica diferenciada del experimento original variando los enfoques considerados y el tamaño y composición de las unidades experimentales, comparando técnicas de desarrollo basadas en SBVR y ODM. Los resultados obtenidos permiten sostener la factibilidad técnica de los mapeos SBVR a OWL 2, como una técnica de desarrollo de ontologías por ingenieros de software.

Palabras-clave: evaluación empírica; experimento; desarrollo de ontologías; mapeo; regla de negocio

Empirical evaluation of business rules mappings for ontology development

Abstract: The methodologies for the development of ontologies aimed to be used as software artifacts remain to be an open research area. Although the most widely recognized are rooted on UML given the acceptance of that language in the community of software engineering, none of them are fully mature. Recent works propose the mapping of SBVR expressions to ontology statements as a building technique, but just an empirical study comparing SBVR and UML-rooted techniques has been performed until now. This work analyzes the technical feasibility of SBVR to OWL 2 mappings by depicting a first differentiated replication of the original experiment varying the treatments, the size and composition of the experimental units, and comparing the performance of SBVR and ODM-based techniques. The findings allow sustaining the technical feasibility of the SBVR to OWL 2 mappings as an ontology development technique able to be applied by the software engineering community.

Keywords: empirical evaluation; experiment; ontology development; mappings, business rule

1. Introducción

Los sistemas de información empresariales podrían evolucionar en versiones más inteligentes mediante la incorporación de las tecnologías semánticas disponibles. En una gran variedad de ejemplos se ha mostrado que la utilización de ontologías - uno de los pilares de las tecnologías semánticas - brindaría importantes beneficios a los sistemas de información empresariales (Ceravolo, Fugazza & Leida, 2007; Karpovic & Nemuraite, 2011; Alberts & Franconi, 2012; Franconi & Mosca, 2012; Beydoun, Low, Tran, & Bogg, 2011; Calero, Ruiz & Piattini, 2006; Myrgioti, Bassiliades & Miliou, 2013; Demuth & Liebau, 2007; Ruotsalo, 2010; Reynares, Caliusco & Galli, 2012; Shue, Chen & Shiue, 2009; Chen, Huang, Bau & Chen, 2012).

Sin embargo, las metodologías para el desarrollo de ontologías aún constituyen un campo abierto de investigación, dado que ninguna de las propuestas actuales se encuentran completamente maduras (Iqbal, Murad, Mustapha, Sharef & Mohd, 2013). La mayoría de las metodologías no proveen suficientes detalles acerca de las técnicas y actividades empleadas, optando por estrategias convencionales para la identificación de los conceptos ontológicos.

Uno de los enfoques más ampliamente conocidos para el desarrollo de ontologías involucra la aplicación de las mejores prácticas del campo de la Ingeniería del Conocimiento. Sin embargo, dichas prácticas usualmente no forman parte del conjunto de herramientas involucradas en el desarrollo de un sistema de información. Su aplicación implica un proceso de aprendizaje dificultoso tanto para profesionales de la ingeniería de software como para investigadores que no se encuentren particularmente familiarizados con tales conocimientos (Gómez-Pérez, Fernández-López & Corcho, 2004).

A fin de evitar tal inconveniente, un grupo de propuestas metodológicas se basó en estándares ampliamente reconocidos en el campo de la ingeniería de software (Nicola, Missikoff & Navigli, 2009). En general, tales propuestas hacen uso del Lenguaje de Modelado Unificado (UML) (Object Management Group (OMG), 2011) y del Lenguaje de Restricción de Objetos (OCL) (Object Management Group (OMG), 2012) para el modelado conceptual de la ontología (Wang y Chan, 2001)(Guizzardi y otros, 2002)(de Nicola y otros, 2009). Sus principales ventajas radican en la amplia aceptación que posee en la comunidad de ingeniería de software, su representación gráfica estandarizada de los modelos, la gran disponibilidad de herramientas que le brindan soporte y su naturaleza extensible. Sin embargo, la falta de una semántica precisa de conjuntos y la ausencia de una teoría de modelos asociada impide la ejecución de razonadores automatizados sobre modelos UML. Tampoco OCL posee una teoría formal de modelos ni una teoría formal de prueba de modelos, por lo que tampoco resulta un candidato posible para la ejecución de procesos automatizados de razonamiento (Object Management Group (OMG), 2009). El lenguaje Ontology Definition Metamodel (ODM) (Object Management Group (OMG), 2009) es otro enfoque en esta misma dirección (Saripalle, Demurjian, Algarín & Blechner, 2013). ODM es una familia de meta-modelos MOF (Object Management Group (OMG),

2011a), mapeos entre dichos meta-modelos, y mapeos desde y hacia el lenguaje UML (Object Management Group (OMG), 2011b), además de un conjunto de perfiles que permiten el modelado de ontologías mediante herramientas basadas en UML.

En forma opuesta a la notación gráfica de las propuestas basadas en UML, trabajos recientes han remarcado que la definición ontológica de conceptos en un lenguaje formal es muy similar a la descripción de dichos conceptos por medio del lenguaje natural: en ambos casos una expresión es construida mediante la combinación de símbolos de acuerdo a reglas gramaticales (Pinker, 2007; Hoekstra, 2009). Basándose en este punto de vista, algunos autores han propuesto el mapeo de expresiones lingüísticas de reglas de negocio a constructores ontológicos como una técnica de desarrollo de ontologías (Ceravolo, Fugazza & Leida, 2007; Demuth & Liebau, 2007; Karpovic & Nemuraite, 2011; Alberts & Franconi, 2012; Franconi & Mosca, 2012; Reynares, Caliusco & Galli, 2013; Reynares, Caliusco & Galli, 2014b).

El lenguaje *Semantics of Business Vocabulary and Business Rules (SBVR)* soporta este enfoque al brindar a los expertos del negocio una herramienta lingüística para especificar semánticamente los conceptos y reglas del negocio en forma independiente del diseño de un sistema de información (Object Management Group (OMG), 2008). La naturaleza lingüística de este lenguaje permite expresar el conocimiento del negocio a través de sentencias en lugar de diagramas. Este enfoque se basa en la idea de que los diagramas son útiles para describir la organización estructural de los conceptos, pero no resultan prácticos a los fines de definir vocabularios y reglas de negocio. SBVR se basa en predicados de lógica de primer orden con algunas extensiones limitadas en lógica modal. Esta fundamentación teórica en lógica formal constituye una característica clave del lenguaje en contextos donde resulta necesario el razonamiento automatizado, presentando una clara ventaja con respecto al uso de modelos UML/OCL.

El objetivo del presente trabajo consiste en evaluar la factibilidad técnica del mapeo de reglas de negocio en sentencias ontológicas para el desarrollo ontológico. Tal objetivo es satisfecho mediante la realización de un experimento exploratorio que compara la performance relativa de métodos basados en los lenguajes UML y SBVR, respectivamente. La comparación se basa en la evaluación de calidad de las ontologías desarrolladas por 10 grupos de igual tamaño conformados por estudiantes avanzados de ingeniería.

2. Diseño del experimento

Este trabajo representa una réplica diferenciada del experimento descrito en Reynares, Caliusco & Galli (2014a), lo cual permite evaluar el grado de generalización de los resultados obtenidos con respecto a la factibilidad técnica de SBVR como lenguaje de modelado para el desarrollo de ontologías.

Mientras un experimento aislado es útil por sí mismo, los estudios replicados resultan claves en el diseño de estudios científicos dado la naturaleza repetible del conocimiento científico. Además, la primera réplica es la más importante dado que establece si una generalización más amplia es posible. La idea de que una réplica implica la repetición exacta del estudio original constituye un error. En la práctica, la repetición de un

experimento debería consistir en llevar a cabo un estudio distinto y evaluar si los resultados se mantienen o no. Esto permite distinguir dos tipos generales de réplicas: (1) las similares y (2) las diferenciadas. Las réplicas similares permiten establecer rápidamente si un resultado se repite o no. En cambio, el objetivo de una réplica diferenciada consiste en extender el rango de condiciones bajo las cuales los resultados se mantienen (Lindsay & Ehrenberg, 1993).

En el experimento descrito en Reynares, Caliusco & Galli (2014a) la hipótesis a evaluar fue:

Hipótesis 1: Las ontologías desarrolladas por medio de un enfoque basado en la utilización del lenguaje SBVR al menos igualan la calidad de las ontologías desarrolladas mediante UML, considerando que las tareas de desarrollo son ejecutadas por recursos humanos con perfil de ingenieros de software sin conocimientos previos en ingeniería ontológica.

UML ha sido la base de diversas propuestas metodológicas para el desarrollo de ontologías; dada su amplia aceptación en la comunidad de ingeniería de software, su representación gráfica estandarizada de los modelos, la gran disponibilidad de herramientas que le brindan soporte y la naturaleza extensible del lenguaje. Pero la falta de una semántica precisa de conjuntos y la ausencia de una teoría de modelos ha dado origen a la definición de ODM, el cual ataca tales inconvenientes mediante un conjunto de meta-modelos formales, perfiles y mapeos.

El experimento descrito en el presente trabajo analiza la factibilidad técnica del mapeo de expresiones de negocios en sentencias ontológicas con el propósito de extender el rango de condiciones bajo las cuales se mantienen los resultados del primer experimento, evaluando la siguiente hipótesis:

Hipótesis 2: Las ontologías desarrolladas por medio de un enfoque basado en la utilización del lenguaje SBVR al menos igualan la calidad de las ontologías desarrolladas mediante ODM, considerando que las tareas de desarrollo son ejecutadas por recursos humanos con perfil de ingenieros de software sin conocimientos previos en ingeniería ontológica.

La hipótesis es evaluada mediante el análisis de calidad de las ontologías desarrolladas por distintos grupos de estudiantes avanzados de ingeniería.

2.1 Enfoques bajo estudio

A continuación se describen las dos propuestas consideradas en el experimento.

1. El desarrollo de ontologías mediante la aplicación de ODM (Object Management Group (OMG), 2009).
2. El desarrollo de ontologías mediante la aplicación del enfoque de mapeo basado en SBVR propuesto en Reynares, Caliusco & Galli (2013, 2014b).

ODM es una familia de meta-modelos MOF (Object Management Group (OMG), 2011a), mapeos entre dichos meta-modelos, y mapeos desde y hacia el lenguaje UML (Object Management Group (OMG), 2011b), además de un conjunto de perfiles que permiten el modelado de ontologías mediante herramientas basadas en UML. Los meta-modelos reflejan la sintaxis abstracta de varios lenguajes estándares utilizados en

el modelado conceptual y la representación de conocimiento. Sin embargo, lo que hace a un buen modelado de software orientado a objetos no necesariamente hace a una buena ontología: una vez que un modelo particular ha sido transformado en una ontología, es necesario tomar precauciones que permitan asegurar que el modelo resultante soportará las aserciones requeridas. A menudo son necesarias reestructuraciones significativas de la ontología para satisfacer tal punto. ODM provee dos maneras de relacionar los modelos a fin de superar ese problema: (1) mediante perfiles UML y (2) mediante mapeos entre los modelos.

El objetivo de un perfil UML es brindar un puente entre UML y las comunidades de representación de conocimiento en una base semántica bien fundamentada, lo cual constituye un medio de relacionar enfoques de software y de modelado lógico para representar conocimiento. ODM define los perfiles para la primera versión de OWL (World Wide Web Consortium (W3C), 2004a), RDFS (World Wide Web Consortium (W3C), 2004b), y Topic Maps (ISO - International Organization for Standardization, 2005b). Aunque los perfiles proporcionan ciertas facilidades para que los usuarios puedan utilizar UML como base para el desarrollo de ontologías para un lenguaje de representación de conocimiento determinado, no facilitan una transformación completa a través del conjunto de paradigmas de representación incluidos en los meta-modelos ODM. Tales necesidades son satisfechas por los mapeos de un meta-modelo a otro.

Por otra parte, el enfoque de mapeo de reglas de negocio propuesto por Reynares, Caliusco y Galli (2013, 2014b) permite la generación automatizable de una ontología OWL 2 (World Wide Web Consortium (W3C), 2009a) mediante la aplicación de un conjunto de transformaciones sobre las especificaciones SBVR de un dominio. SBVR ha sido concebido para expertos del negocio y diseñado para ser utilizado a fines de satisfacer los requerimientos del negocio. La naturaleza lingüística del lenguaje permite la expresión del conocimiento del negocio mediante sentencias en lugar de diagramas, lo cual responde a la idea de que los diagramas permiten describir estructuralmente los conceptos pero no resultan prácticos a los fines de definir vocabularios y expresar reglas de negocio. SBVR está basado en lógica de predicados de primer orden con algunas extensiones restringidas en lógica de alto nivel y lógica modal.

2.2 Contexto y unidades experimentales

El experimento fue realizado en el contexto del curso de grado denominado "Desarrollo de Sistemas de Información basados en Ontologías", el cual forma parte del último nivel de la carrera de Ingeniería en Sistemas de Información de la Universidad Tecnológica Nacional en la provincia de Santa Fe, Argentina. Las unidades experimentales consistieron en 10 grupos de igual tamaño, conformados por 20 estudiantes del curso. Un proceso aleatorio fue utilizado para la generación de muestras de igual tamaño y la asignación de las unidades experimentales a los distintos enfoques. De esta forma, 5 grupos aplicaron el enfoque de mapeo basado en SBVR, mientras los restantes 5 siguieron la propuesta ODM. Los participantes obtuvieron créditos académicos por formar parte del experimento. Además, respondieron una encuesta anónima acerca de su conocimiento previo en relación a los temas involucrados en el experimento, evaluando las mismas en un rango de valores de 1 - conocimientos previos nulos - a 10 - conocimientos previos de nivel profesional -.

Aunque la naturaleza anónima de la encuesta buscaba obtener respuestas sinceras, es necesario considerar el sesgo subjetivo presente en cualquier auto evaluación. En las Figuras 1-3 se muestran las comparaciones de frecuencia sobre conocimientos previos en (P1) el uso de artefactos UML, (P2) el desarrollo de sentencias lógicas y (P3) la ingeniería ontológica, respectivamente¹. Para evaluar las diferencias en las respuestas de los enfoques comparados fue utilizado el test de Mann-Whitney-Wilcoxon (MWW). Este test (también denominado estadístico U o simplemente U) es un test no paramétrico de la hipótesis nula - donde dos poblaciones resultan iguales - contra una hipótesis alternativa - donde una población particular tiende a tener valores mayores que la otra -. El test posee una eficiencia mayor al test "t" en distribuciones no-normales y una eficiencia cercana a la del test "t" en distribuciones normales (Mann & Whitney, 1947; Fay & Proschan, 2010).

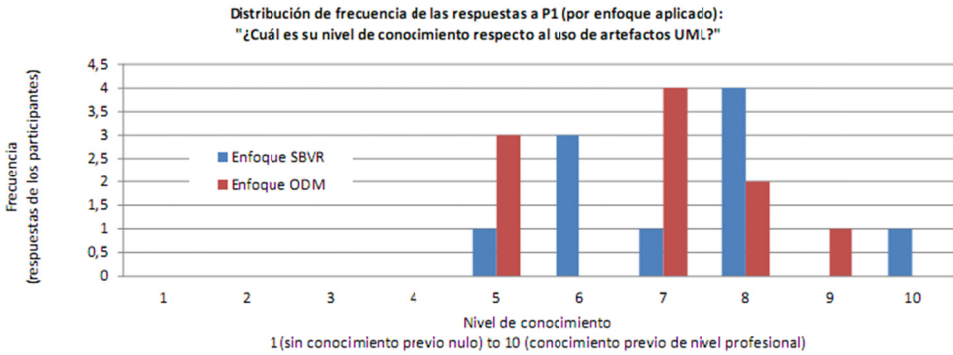


Figura 1 - Distribución de frecuencia de las respuestas a P1 (por enfoque aplicado): "¿Cuál es su nivel de conocimiento respecto al uso de artefactos UML?"

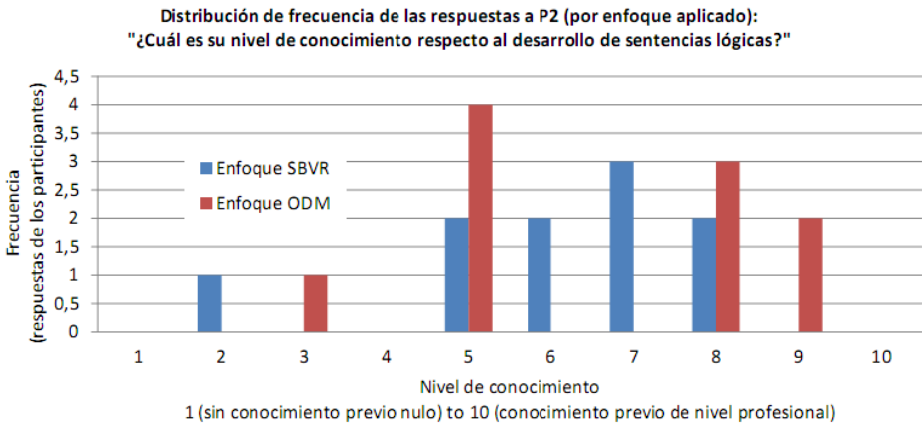


Figura 2 - Distribución de frecuencia de las respuestas a P2 (por enfoque aplicado): "¿Cuál es su nivel de conocimiento respecto al desarrollo de sentencias lógicas?"

¹ La encuesta completa y sus resultados pueden encontrarse en <https://code.google.com/p/development-of-ontologies-experiment-replicated/>

La Tabla 1 muestra los intervalos críticos de U para dos muestras de igual tamaño A y B (donde el tamaño $n = 10$), para test direccionales y no direccionales y para los niveles de significancia más comúnmente utilizados. La hipótesis nula- ambos poblaciones tienden a poseer los mismo valores - se acepta si los valores observados de U se encuentran en el rango comprendido entre los límites superiores e inferiores. La Tabla 2 presenta el valor observado de U para ambos enfoques en cada una de las preguntas de la encuesta. Los valores observados permiten concluir que no existe diferencia estadísticamente significativa entre ambos enfoques en relación a su conocimiento previo sobre P1, P2, y P3. Tales resultados permitieron concluir que ambos grupos poseían el mismo nivel de conocimientos previos en relación a los temas involucrados en el experimento, los cuales se asemejan al conocimiento que un ingeniero de software sin formación previa en ingeniería ontológica posee inicialmente.

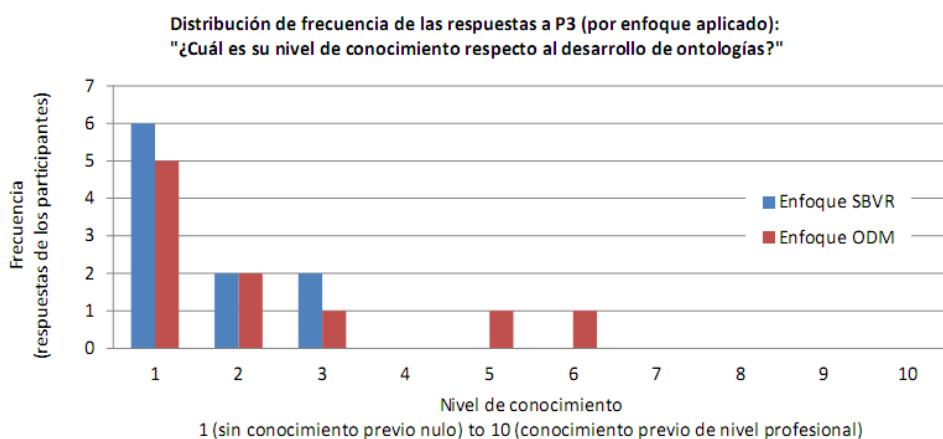


Figura 3 - Distribución de frecuencia de las respuestas a P2 (por enfoque aplicado): "¿Cuál es su nivel de conocimiento respecto al desarrollo de ontologías?"

Tabla 1 – Intervalos críticos de U para dos muestras de tamaño $n = 10$

Intervalos críticos de U			
	Nivel de significancia para:		
	Tests Direccionales		
	0.05	0.025	0.01
	Tests no-Direccionales		
	-	0.05	0.02
Límite inferior	27	23	19
Límite superior	73	77	81

Tabla 2 – Valores de U por enfoque aplicado

Valores de U		
	Enfoque SBVR	Enfoque ODM
<i>Conocimientos previos en:</i>		
<i>(P1) el uso de artefactos UML</i>	42.5	57.5
<i>(P2) el desarrollo de sentencias lógicas</i>	56	44
<i>(P3) ingeniería ontológica</i>	58	42

2.3 Tarea y materiales

La tarea del experimento consistió en la especificación ontológica de las políticas de asignación de becas estudiantiles de la universidad, especificadas originalmente en un documento oficial de la institución y descritas en lenguaje natural. Con respecto a las herramientas, el enfoque basado en SBVR utilizó editores de texto convencionales para modelar el dominio del negocio, mientras aquellos que aplicaron ODM utilizaron la herramienta Microsoft Visio para el desarrollo de los modelos gráficos. Ambos enfoques implementaron la ontología en OWL 2 por medio de Protégé, un editor libre y de código abierto.

3. Resultados

La performance de los dos enfoques fue comparada mediante la evaluación de la calidad de las ontologías desarrolladas. Tal evaluación fue realizada por los autores de este trabajo mediante OQuaRE (Duque-Ramos, Fernández-Breis, Iniesta, Dumontier, Aranguren, Schulz, Aussenac-Gilles & Stevens, 2013; Duque-Ramos, Fernández-Breis, Stevens & Aussenac-Gilles, 2011), un marco de trabajo basado en el estándar SQuaRE para la evaluación de la calidad del software (International Organization for Standardization (ISO), 2005). OQuaRE evalúa las ontologías independientemente de cualquier proceso de desarrollo particular, lo cual consituye un proceso objetivo de evaluación capaz de ser replicado.

OQuaRE define un modelo de calidad y sus respectivas métricas para la evaluación de las ontologías. El modelo de calidad es dividido en dimensiones - o características - organizadas en sub-dimensiones - o sub-características -, las cuales son evaluadas por medio de un conjunto de métricas. OQuaRE define el criterio para normalizar los valores de las métricas en un rango de 1 a 5: calidad no aceptable es asociada al valor 1, la mínima calidad aceptable responde al valor 3, y el valor 5 excede los requerimientos de calidad. Luego de la normalización, el puntaje asociado a cada una de las características es la media de las sub-características asociadas, cuyo valor es calculado como la media de sus métricas. El resultado final de la evaluación de calidad consiste en el conjunto de valores para cada una de las características evaluadas. Esto permite la identificación de las fortalezas y debilidades de las ontologías consideradas en lugar de señalar la "mejor" ontología, como se propone en los trabajos de Lozano-Tello & Gómez-Pérez (2004); Park, Oh & Ahn (2011); y Vrandecic (2010).

Las dimensiones de calidad evaluadas en el presente experimento se describen a continuación:

- La *dimensión estructural* involucra propiedades formales y semánticas que resultan importantes al evaluar ontologías, tales como consistencia, nivel de formalización, redundancia o complejidad.
- La *dimensión de adecuación funcional* refiere al grado de alineamiento de la ontología para su finalidad prevista, de acuerdo a las categorías identificadas por Stevens, Wroe, Gobel & Lord, 2008.
- La *dimensión de mantenibilidad* está relacionada a la capacidad de las ontologías de ser modificadas de acuerdo a cambios en el entorno, en los requerimientos o en las especificaciones funcionales.
- La *dimensión de compatibilidad* refiere a la habilidad de dos o más ontologías para intercambiar información y/o cumplir su función mientras comparten el mismo entorno de hardware o software. Esta dimensión puede ser evaluada sobre una única ontología - aunque intuitivamente involucra las propiedades de más de una -, dado que es cuantitativamente evaluada por medio de un conjunto de métricas aplicadas a cada ontología en forma separada.
- La *dimensión de transferibilidad* es el grado con el cual la ontología puede ser transferida de un entorno a otro. La *dimensión de operatividad* refiere al esfuerzo necesario para utilizar la ontología por un determinado conjunto de usuarios.
- La *dimensión de confiabilidad* es la capacidad de la ontología de mantener su nivel de performance bajo condiciones establecidas, por un período dado de tiempo.

Tres métricas no fueron consideradas en la evaluación de las dimensiones de calidad. Las métricas denominadas *riqueza de las anotaciones* - número medio de anotaciones por clase - y *riqueza de clases* - número medio de instancias por clase - no fueron evaluadas dado que la anotación e instanciación de ontologías no formaba parte del experimento. La métrica denominada *riqueza de atributos* - número medio de atributos por clase - no fue evaluada dado que los atributos no forman parte de la especificación estructural de lenguaje de implementación de las ontologías (World Wide Web Consortium (W3C), 2009b).

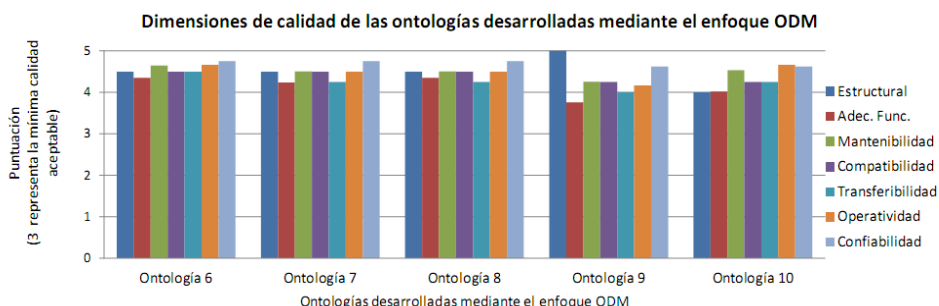


Figura 4 - Puntuación de las dimensionaes de calidad de las ontologías desarrolladas mediante el enfoque SBVR

OQuaRE define también las dimensiones denominadas *eficiencia de performance* y *calidad de uso*. La eficiencia de performance expone la relación entre el nivel de performance de la ontología y los recursos utilizados bajo condiciones establecidas, tomando en cuenta elementos tales como tiempo de respuesta o consumo de memoria. La calidad en uso refiere al grado con el cual la ontología satisface los requerimientos de usuarios específicos. Sin embargo, tales dimensiones no fueron consideradas dado la inexistencia de métricas asociadas a sus correspondientes sub-características. La Figura 4 muestra los niveles de calidad de las ontologías desarrolladas por medio del enfoque basado en el mapeo de expresiones SBVR.

La Figura 5 muestra los niveles de calidad de las ontologías desarrolladas mediante el estándar ODM.

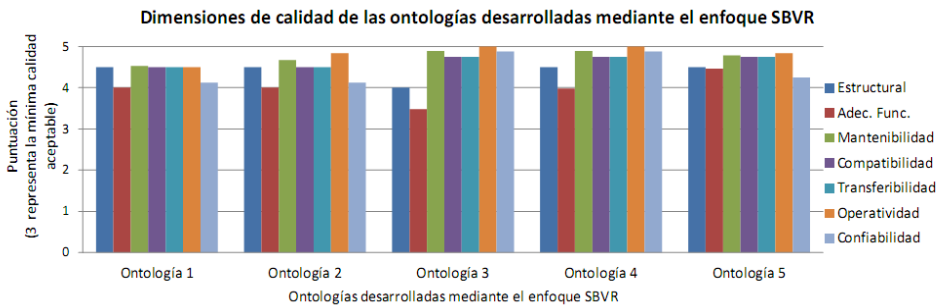


Figura 5 - Puntuación de las dimensiones de calidad de las ontologías desarrolladas mediante el enfoque ODM

5. Análisis de resultados

Un rápido vistazo a las Figuras 4 y 5 permite reconocer un primer resultado importante: de acuerdo a OQuaRE, todas las ontologías superan la mínima calidad aceptable. La Tabla 3 muestra el valor medio de cada una de las dimensiones de calidad evaluadas: la primer columna muestra el nivel de calidad de las ontologías desarrolladas mediante el enfoque basado en SBVR mientras la segunda presenta los valores de las ontologías obtenidas por medio de ODM. Además, la última fila de la tabla muestra el valor medio de la calidad de la ontología de acuerdo al enfoque seguido. Una comparación fila por fila permite observar que las medias de las dimensiones de calidad son muy similares entre ambos enfoques, incluso en sus valores medios globales. Nuevamente, el test MWW permitió evaluar las diferencias entre los valores obtenidos. La Tabla 4 muestra los intervalos críticos de U para dos muestras de tamaño $n = 5$, tanto para test direccionales como no-direccionales y para los niveles de significancia más comúnmente considerados. Finalmente, la Tabla 5 muestra los valores de U para cada enfoque en cada una de las dimensiones de calidad evaluadas. Los valores observados permiten concluir que no existen diferencias estadísticamente significativas entre ambos enfoques, en ninguna de las dimensiones evaluadas.

Tabla 3 – Valores medios por dimensión de calidad y enfoque aplicado

Valores medios por dimensión de calidad		
	Enfoque SBVR	Enfoque ODM
<i>Estructural</i>	4.40	4.50
<i>Adecuación Funcional</i>	3.99	4.14
<i>Mantenibilidad</i>	4.75	4.49
<i>Compatibilidad</i>	4.65	4.40
<i>Transferibilidad</i>	4.65	4.25
<i>Operatividad</i>	4.83	4.50
<i>Confiabilidad</i>	4.45	4.70
<i>Media Global</i>	4.53	4.43

Tabla 4 – Intervalos críticos de U para dos muestras de tamaño n = 5

Intervalos críticos de U			
Nivel de significancia para:			
Tests Direccionales			
	0.05	0.025	0.01
Tests no-Direccionales			
	-	0.05	0.02
<i>Límite inferior</i>	4	2	1
<i>Límite superior</i>	21	23	24

6. Discusión

El experimento presentado se encuentra sesgado en forma negativa con respecto al enfoque basado en el mapeo de expresiones SBVR, dado que los participantes poseían conocimientos previos en el uso de artefactos UML mientras el conocimiento acerca del lenguaje SBVR era inexistente.

Resulta interesante destacar el potencial del lenguaje SBVR para expresar nociones complejas de un dominio de interés: mientras las ontologías desarrollas aplicando el enfoque basado en mapeos aprovechan la totalidad del poder expresivo del lenguaje OWL, las desarrolladas mediante ODM sólo comprenden los aspectos básicos de OWL. Aunque no se vea reflejado en los resultados formales del experimento, un aspecto importante de resaltar lo constituyen las sensaciones de los participantes del experimento. Aquellos que aplicaron el enfoque SBVR manifestaron la *simpleza* del proceso de desarrollo basado en la descripción de un dominio por medio del lenguaje natural. En cambio, aquellos que utilizaron ODM expresaron su descontento con un proceso que consideraron dificultoso de ejecutar y cuyos resultados mostraban una complejidad y dimensión que excedía la del dominio real que era objeto del modelado.

La Tabla 6 muestra el poder expresivo de cada una de las ontologías mientras la Tabla 7 presenta los constructores de la Lógica Descriptiva y sus nombres asociados.

Las ontologías desarrolladas fueron examinadas por un experto del dominio para su posterior evaluación de calidad, a fin de obtener cierta retroalimentación respecto a la brecha semántica existente entre la realidad y el modelado realizado. Más allá de las diferencias esperables entre cada modelado de la misma realidad, la totalidad de las ontologías representaban el dominio en forma adecuada. Sin embargo, una evaluación más sistemática y compleja eliminaría cualquier sesgo subjetivo.

Tabla 5 – Valores de U por dimensión de calidad y enfoque aplicado

Valores de U por dimensión de calidad		
	<i>Enfoque SBVR</i>	<i>Enfoque ODM</i>
<i>Estructural</i>	14.5	10.5
<i>Adecuación Funcional</i>	17	8
<i>Mantenibilidad</i>	2	23
<i>Compatibilidad</i>	11.5	13.5
<i>Transferibilidad</i>	3	22
<i>Operatividad</i>	3	22
<i>Confiabilidad</i>	15	10

Tabla 6 – Expresividad de la Lógica Descriptiva por ontología y por enfoque aplicado

Expresividad DL		
<i>Enfoque SBVR</i>	Ontología 1	ALCHIC(D))
	Ontología 2	ALCIQ(D)
	Ontología 3	ALCRQ(D)
	Ontología 4	ALIQ(D)
	Ontología 5	ALCROIQ(D)
<i>Enfoque ODM</i>	Ontología 6	ALU(D)
	Ontología 7	ALUI(D)
	Ontología 8	ALCF(D)
	Ontología 9	ALCO(D)
	Ontología 10	ALCIQ(D)

Aunque cualquier experimento de laboratorio sufre una cierta falta de realismo, un estudio de campo de estas características resulta complejo de diseñar y ejecutar. En cuanto a los sujetos experimentales, existen buenas razones para realizar experimentos con estudiantes. Por ejemplo, a fin de evaluar diseños experimentales y posibilitar un rápido descubrimiento de resultados negativos en los cuales no tenga sentido

profundizar (Tichy, 2000). Los estudiantes también resultan representativos de profesionales de escasa experiencia (Sjøberg, Hannay, Hansen, Kampenes, Karahasanovic, Liborg, & Rekdal, 2005), lo cual es actualmente la condición general de los profesionales de ingeniería de software en relación al uso de tecnologías semánticas.

Tabla 7 – Constructores DL y nombres de lenguaje: "A" refiere a conceptos atómicos, "C" y "D" a cualquier definición de concepto, "R" a roles atómicos y "S" a definición de roles

Constructores DL		
Constructor	Sintaxis	Lenguaje
Concepto	A	FL ₀
Nombre de rol	R	
Intersección	$C \cap D$	
Restricción de valor	$\forall R.C$	FL-
Cuantificación existencial limitada	$\exists R$	AL
Universal		S
Vacío	\perp	
Negación atómica	$\neg A$	
Negación	$\neg C$	C
Unión	$C \cup D$	U
Restricción existencial	$\exists R.C$	E
Restricción numérica	$(\geq nR) (\leq nR)$	N
Nominales	$\{a_1 \dots a_n\}$	O
Jerarquía de rol	$(R \subseteq S)$	H
Rol inverso	R^-	I
Restricción cuantificada numéricamente	$(\geq nR.C) (\leq nR.C)$	Q

7. Conclusiones y trabajos futuros

Un requerimiento clave para la adopción de las tecnologías semánticas en entornos reales es la disponibilidad de tecnologías probadas, las cuales garanticen una ingeniería eficiente de ontologías de alta calidad. Sin embargo, tales garantías sólo pueden ser provistas mediante un estudio empírico de las metodologías, técnicas y enfoques propuestos en los ámbitos académicos.

Este trabajo representa una primera réplica diferenciada de un experimento previo, evaluando el grado de generalización de los resultados originales y extendiendo el rango de condiciones bajo las cuales tales resultados se mantienen. Los resultados obtenidos en la réplica presentada permiten sostener aquellos obtenidos en el estudio original. La factibilidad técnica de mapear expresiones SBVR en ontologías OWL 2 ha

sido demostrada, incluso variando el tamaño y composición de las unidades experimentales y los enfoques bajo estudio. Tales resultados refuerzan el potencial de los mapeos SBVR a OWL 2 como una técnica de desarrollo de ontologías digno de un estudio más profundo.

De acuerdo a estos resultados, los trabajos futuros involucran tres aspectos centrales. El primero consiste en la validación, formalización y extensión de los mapeos SBVR a OWL 2 propuestos actualmente. Luego, resulta necesaria la definición de mapeos de SBVR a Cláusulas de Horn expresadas en lenguaje SWRL, estrechando de esta manera la brecha existente en el poder expresivo de SBVR y OWL 2. Finalmente, y desde un punto de vista empírico, el próximo paso consiste en la realización de un experimento controlado destinado a aislar las relaciones entre las dimensiones de calidad de las ontologías y los métodos/prácticas ejecutadas en su desarrollo. Dicho experimento requiere de un diseño más complejo y detallado a fin de balancear exitosamente los aspectos de monitoreo y realismo, a fin de ser capaces de generalizar los resultados a entornos industriales. Además, la evaluación de las ontologías por medio de diversos marcos de trabajo proveería un panorama más completo de su nivel de calidad.

Referencias bibliográficas

- Alberts, R. & Franconi, E. (2012). An integrated method using conceptual modelling to generate an ontology-based query mechanism. In P. Klinov, & M. Horridge (Eds.) *OWL: Experiences and Directions Workshop 2012 (OWLED)*. Greece volume 849 of CEUR Workshop Proceedings.
- Beydoun, G., Low, G., Tran, N. & Bogg, P. (2011). Development of a peer-to-peer information sharing system using ontologies. *Expert Systems with Applications*, 38(8), 9352-9364. doi: 10.1016/j.eswa.2011.01.104
- Calero, C., Ruiz, F. & Piattini, M. (2006). *Ontologies for Software Engineering and Software Technology*. Springer, Heidelberg.
- Ceravolo, P., Fugazza, C. & Leida, M. (2007). Modeling semantics of business rules. In *EcoSystems, Digital Technologies Conference. DEST 07*. 171-176.
- Chen, R.C., Huang, Y.H., Bau, C.T. & Chen, S.M. (2012). A recommendation system based on domain ontology and SWRL for anti-diabetic drugs selection. *Expert Systems with Applications*, 39, 3995-4006. doi: 10.1016/j.eswa.2011.09.061
- Demuth, B. & Liebau, H.B. (2007). An approach for bridging the gap between business rules and the semantic web. In A. Paschke & Y. Biletskiy (Eds.), *Proceedings of the International Symposium on Advances in Rule Interchange and Applications. LNCS, 4824*, 119-133.
- Duque-Ramos, A., Fernández-Breis, J. T., Iniesta, M., Dumontier, M., Aranguren, M. E., Schulz, S., Aussenac-Gilles, N. & Stevens, R. (2013). Evaluation of the OQuaRE framework for ontology quality. *Expert Systems with Applications*, 40(7), 2696-2703. doi: 10.1016/j.eswa.2012.11.004

- Duque-Ramos, A., Fernández-Breis, J. T., Stevens, R. & Aussenac-Gilles, N. (2011). OQuaRE: A SQuaRE-based approach for evaluating the quality of ontologies. *Journal of Research and Practice in Information Technology*, 43(2), 159-176.
- Fay, M. & Proschan, M. (2010). Wilcoxon Mann Whitney or t-test? On assumptions for hypothesis tests and multiple interpretations of decision rules. *Statistics Surveys*, 4, 1-39. doi: 10.1214/09-SS051
- Franconi, E. & Mosca, A. (2012). The formalisation of ORM2 and its encoding in OWL2 . *Technical Report KRDB Research Centre Technical Report KRDB 12-2*, Faculty of Computer Science, Free University of Bozen-Bolzano Italy.
- Gómez-Pérez, A., Fernández-López, M. & Corcho, O. (2004). *Ontological Engineering*. Springer/Heidelberg.
- Hoekstra, R. (2009). *Ontology Representation - Design Patterns and Ontologies that Make Sense. Frontiers in Artificial Intelligence and Applications*, 197. IOS Press.
- International Organization for Standardization (ISO) (2005). ISO/IEC 25000:2005, Software Engineering - Software Product Quality Requirements and Evaluation (SQuaRE).
- International Organization for Standardization (ISO) (2005b). ISO/IEC 13250-2: Topic Maps Data Model.
- Iqbal, R., Murad, M. A. A., Mustapha, A., Sharef & Mohd, N. (2013). An analysis of ontology engineering methodologies: A literature review. *Research Journal of Applied Sciences, Engineering and Technology*, 6, 2993-3000.
- Karpovic, J. & Nemuraite, L. (2011). Transforming SBVR business semantics into Web Ontology Language OWL 2: Main concepts. In *Proceedings of 17th International Conference on Information and Software Technologies*
- Lindsay, R. M. & Ehrenberg, A. S. (1993). The design of replicated studies. *The American Statistician*, 47, 217-228.
- Lozano-Tello, A. & Gómez-Pérez, A. (2004). Ontometric: A method to choose the appropriate ontology. *Journal of Database Management*, 15(2), 1–18. doi: 10.4018/jdm.2004040101
- Mann, H. B. & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *Annals of Mathematical Statistics*, 18, 50-60.
- Myrgioti, E., Bassiliades, N. & Miliou, A. (2013). Bridging the HASM: An OWL ontology for modeling the information pathways in haptic interfaces software. *Expert Systems with Applications*, 40(4), 1358-1371. doi: 10.1016/j.eswa.2012.08.053
- Nicola, A. D., Missikoff, M. & Navigli, R. (2009). A software engineering approach to ontology building. *Information Systems*, 34(2), 258-275. Doi: 10.1016/j.is.2008.07.002

- Object Management Group (OMG) (2008). Semantics of Business Vocabulary and Business Rules (SBVR). Version 1.0: Formal Specification.
- Object Management Group (OMG) (2009). Ontology Definition Metamodel (ODM). V. 1.0.
- Object Management Group (OMG) (2011a). Meta Object Facility (MOF). V. 2.4.1.
- Object Management Group (OMG) (2011b). Unified Modeling Language (UML). V. 2.4.1.
- Park, J., Oh, S. & Ahn, J. (2011). Ontology selection ranking model for knowledge reuse. *Expert Systems with Applications*, 38(5), 5133–5144. doi: 10.1016/j.eswa.2010.10.002
- Pinker, S. (2007). *The Stuff of Thought*. London, England. Allen Lane.
- Reynares, E., Caliusco, M. L. & Galli, M. R. (2012). EDON: A method for building an ontology as software artefact. In *41st Argentine Conference on Informatics - 13th Argentine Symposium on Software Engineering*. Bs As, Argentina.
- Reynares, E., Caliusco, M. L. & Galli, M. R. (2013). An automatable approach for SBVR to OWL 2 mappings. In *XVI Ibero-American Conference on Software Engineering (CIbSE 2013)*. Montevideo, Uruguay.
- Reynares, E., Caliusco, M. L. & Galli, M. R. (2014a). Approaching the feasibility of SBVR as modeling language for ontology development: An exploratory experiment. *Expert Systems with Applications*, 41(2), 1576–1583. doi: 10.1016/j.eswa.2013.08.054
- Reynares, E., Caliusco, M. L. & Galli, M. R. (2014b). SBVR to OWL 2 mappings: An automatable and structural-rooted approach. *CLEI-EJ*, 17.
- Ruotsalo, T. (2010). *Methods and Applications for Ontology-Based Recommender Systems (PhD. Thesis)*. Ph.D Thesis Alto University School of Science and Technology Finland.
- Saripalle, R. K., Demurjian, S. A., Algarín, A. D. L. R. & Blechner, M. (2013). A software modeling approach to ontology design via extensions to ODM and OWL. *International Journal on Semantic Web and Information Systems (IJSWIS)*, 9(2), 62–97. doi: 10.4018/jswis.2013040103
- Shue, L.Y., Chen, C.W. & Shiue, W. (2009). The development of an ontology-based expert system for corporate financial rating. *Expert Systems with Applications*, 36(2), 2130–2142. doi: 10.1016/j.eswa.2007.12.044
- Sjøberg, D. I., Hannay, J. E., Hansen, O., Kampenes, V. B., Karahasanovic, A., Liborg, N.K., & Rekdal, A. C. (2005). A survey of controlled experiments in software engineering. *IEEE Transactions on Software Engineering*, 31(9), 733–753. doi: 10.1109/TSE.2005.97
- Stevens, R., Wroe, C., Gobel, C. & Lord, P. (2008). Application of ontologies in bioinformatics. In S. Staab, & R. Studer (Eds.), *Handbook of Ontologies in Informations Systems*. 635–658. Springer.

- Tichy, W. F. (2000). Hints for reviewing empirical work in software engineering. *Empirical Software Engineering*, 5(4), 309-312. doi: 10.1023/A:1009844119158
- Vrandecic, D. (2010). Ontology Evaluation. Ph.D thesis. University of Karlsruhe.
- World Wide Web Consortium (W3C) (2004a). OWL Web Ontology Language Overview.
- World Wide Web Consortium (W3C) (2004b). RDF Vocabulary Description Language 1.0: RDF Schema.
- World Wide Web Consortium (W3C) (2009a). OWL 2 Web Ontology Language. Document Overview.
- World Wide Web Consortium (W3C) (2009b). OWL 2 Web Ontology Language. Structural Specification and Functional-Style Syntax.

