

# Metodologías y Herramientas para el Diseño Conceptual con el Soporte de Razonamiento

Germán Braun<sup>1,2,3</sup>

Laura Cecchi<sup>1</sup>

Pablo Fillottrani<sup>2,4</sup>

email: {german.braun,lcecchi}@fi.uncoma.edu.ar, prf@cs.uns.edu.ar

<sup>1</sup>*Grupo de Investigación en Lenguajes e Inteligencia Artificial*

Departamento de Teoría de la Computación - Facultad de Informática

UNIVERSIDAD NACIONAL DEL COMAHUE

<sup>2</sup>*Laboratorio de I&D en Ingeniería de Software y Sistemas de Información*

Departamento de Ciencias e Ingeniería de la Computación

UNIVERSIDAD NACIONAL DEL SUR

<sup>3</sup>*Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)*

<sup>4</sup>*Comisión de Investigaciones Científicas de la provincia de Buenos Aires (CIC)*

## Resumen

Esta línea de investigación se desarrolla en forma colaborativa entre docentes-investigadores de la Universidad Nacional del Comahue y de la Universidad Nacional del Sur, en el marco de proyectos de investigación financiados por las universidades antes mencionadas.

El objetivo general del trabajo de investigación es definir metodologías para el soporte a la ingeniería ontológica. El énfasis está puesto en integrar las características intuitivas inherentes al modelado conceptual con los sistemas lógicos y gráficos en ambientes de desarrollo. Se pretende trabajar en algoritmos de *layout* gráfico de ontologías, incorporación de resultados deducibles y su adaptación a nuevas arquitecturas de sistemas. De esta forma, se podrán visualizar diagramas obtenidos a partir de deducciones, cuya complejidad dependerá de la expresividad del lenguaje y del sistema de razonamiento subyacente.

**Palabras Clave:** Ingeniería de Software basada en Conocimiento, Ingeniería de Conocimiento, Lógicas Descriptivas, Ontologías.

## Contexto

Este trabajo está parcialmente financiado por la Universidad Nacional del Comahue, en el marco del proyecto de investigación *Agentes Inteligentes en Ambientes Dinámicos (04/F006)*, por la Universidad Nacional del Sur a través del proyecto de investigación *Integración de Información y Servicios en la Web (24/N027)* y por el Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), en el contexto de una beca interna doctoral. Los proyectos de investigación tienen una duración de cuatro años y la beca doctoral una duración de 5 años, finalizando esta última en abril de 2019.

## 1. Introducción

La integración de la información es uno de los problemas centrales que se presentan en los sistemas de información cooperativos [3]. Esta integración requiere de dos factores fundamentales para el diseño y mantenimiento de aplicaciones como son el modelado conceptual del dominio y el soporte de razonamiento sobre la representación conceptual.

Este nivel conceptual contiene una descripción formal de los conceptos y las relaciones entre ellos. La característica principal es que tal descripción es independiente de cualquier sistema y está orientada a expresar la semántica de las aplicaciones. En particular, estos modelos deberían ser capaces de representar relaciones entre datos que provengan de diferentes fuentes.

Las ontologías conforman el esqueleto fundamental del concepto de la Web Semántica [1]. Pero su desarrollo, mantenimiento, y validación no es una tarea fácil. Se deben tener en cuenta el conocimiento de los expertos, los sistemas de información preexistentes y los nuevos requerimientos para poder obtener ontologías de calidad que sean útiles en la Web Semántica. Esta ingeniería ontológica necesita de metodologías y herramientas para la creación, edición y actualización de ontologías que permitan establecer criterios de calidad claros y medibles. Existen varias herramientas entre las que podemos nombrar a NeOn toolkit [11], TopBrain Composer [17] y Kaon [13]. Una de las herramientas más populares es Protégé [12] desarrollada por la Universidad de Stanford. Protégé ofrece un amplio conjunto de estructuras para modelar conocimiento que soportan la creación, visualización y manipulación de ontologías. Permite modelar ontologías mediante los editores Protege-Frames y Protégé-OWL. Este último provee una herramienta visual para mostrar las inferencias obtenidas de razonadores OWL [14] externos, aunque sólo limitado a relaciones IsA. Sin embargo, el constante avance de los lenguajes ontológicos y sistemas de razonamiento, requiere que estas herramientas presenten nuevas características que den soporte a la manipulación de ontologías cada vez más expresivas. En esta dirección, surge ICOM [7] como una nueva herramienta para diseñar, mantener y vincular múltiples ontologías, facilitando su evolución en el tiempo. ICOM emplea razonamiento lógico completo para verificar especificaciones, inferir hechos implícitos y manifestar inconsistencias. Además, promueve la combinación del diseño gráfico de ontologías

con el soporte para razonamiento, la cual es una característica faltante en las herramientas previamente mencionadas.

En el ámbito de esta investigación, nosotros proponemos ampliar el soporte para el modelado ontológico mediante metodologías que integren el soporte gráfico con las necesidades de los usuarios y el razonamiento automático. El objetivo es que los usuarios puedan visualizar las ontologías durante su evolución junto con todas las deducciones expresadas en ese mismo lenguaje gráfico.

La estructura del presente trabajo es la siguiente. En la sección 2 presentamos los objetivos de los proyectos de investigación en los que se enmarca este trabajo y describimos brevemente la línea de investigación. En la sección 3 se introduce una descripción del problema que se estudia, la motivación y los objetivos de la investigación. En la sección 4 indicamos algunos resultados obtenidos y trabajos futuros. Finalmente, comentamos aspectos referentes a la formación de recursos humanos en esta temática.

## 2. Línea de Investigación y Desarrollo

El proyecto de investigación *Agentes Inteligentes en Ambientes Dinámicos* tiene varios objetivos generales. Uno de ellos es el de *desarrollar conocimiento especializado en el área de Inteligencia Artificial*. En este sentido, se estudian técnicas de representación de conocimiento y razonamiento, métodos de planificación y tecnologías del lenguaje natural aplicadas al desarrollo de sistemas multiagentes.

Por otro lado, en el proyecto de investigación *Integración de Información y Servicios en la Web* se propone investigar y desarrollar metodologías y herramientas que favorezcan la interoperabilidad semántica de información y de servicios en la Web, fundamentados en los últimos avances en el área de lenguajes de representación del conocimiento, ontologías y modelado conceptual.



Ambos proyectos confluyen en la línea de investigación de este trabajo, en la que se explora entre otros, sobre temas afines a la Representación del Conocimiento, las Lógicas Descriptivas [4], las Ontologías, la Ingeniería de Software basada en Conocimiento y la Ingeniería de Conocimiento. Particularmente, se ha escogido experimentar sobre metodologías que integren razonamiento con un *front-end* gráfico para dar soporte a la ingeniería de ontologías.

### 3. Metodologías basadas en la integración de Soporte Gráfico y Razonamiento para Modelado Conceptual

El modelado conceptual permite describir la semántica de las aplicaciones de software en un alto nivel de abstracción. Su principal objetivo es identificar, analizar y describir los conceptos esenciales y restricciones de un universo de interés con la ayuda de un lenguaje de modelado que está basado en un conjunto de conceptos de modelado básicos. Ejemplos de tales lenguajes son ER [5], ORM [10] y UML [2]. En este contexto, en [8] Guizzardi considera a una ontología como un tipo especial de especificación, por lo tanto, es posible considerar al modelado ontológico como un tipo especial de modelado conceptual.

Una ontología es un “acuerdo” sobre una representación formal, explícita, compartida y parcial de una conceptualización [16]. La característica formal dada a las ontologías requiere que éstas sean entendibles por los sistemas computacionales, y a diferencia de los modelos de datos, su valor fundamental es su relativa independencia de aplicaciones particulares. Todos estos requerimientos hacen que el desarrollo y mantenimiento de ontologías sean tareas altamente complejas dentro de la ingeniería de ontologías. Por lo tanto, es necesario dar soporte a los usuarios facilitando ambientes de desarrollo que inte-

gren la semántica formal de los modelos con las intenciones de los modeladores. La representación gráfica junto con los sistemas de razonamiento son esenciales para la gestión del conocimiento tanto explícito como implícito.

Todos los sistemas lógicos suponen objetos ideales pero ellos no consideran la información de la realidad o la intuición. Esto explica que sólo aquellos hechos implicados por el modelo actual serán sugeridos como posibles evoluciones. Sin embargo, si nosotros pensamos al modelado como un enfoque basado en hechos empíricos, vamos a necesitar algo más que sólo la lógica formal. Son necesarias observaciones, experimentos y análisis de patrones para poder confirmar nuestras hipótesis.

Si bien los modelos conceptuales que permiten reusar y compartir conocimiento, además de posibilitar razonamiento sobre el modelo, han ido progresando, este incremento en el desarrollo de ontologías no fue controlado. Inclusive, inicialmente, era realizado de una manera *ad hoc* por parte de los expertos del dominio. El resultado eran ontologías de baja calidad debido al uso limitado de los lenguajes en los cuales eran escritas, y a la imposibilidad, de los expertos, de capturar el conocimiento del dominio. Ante esta situación, la necesidad de metodologías y herramientas que den soporte a algunas tareas relacionadas a la ingeniería ontológica fue observada y relevada en [6] dando lugar a un nuevo paradigma de herramientas de visualización de ontologías.

En este contexto y, dada la disparidad, aún existente, entre la evolución de los lenguajes de ontologías y los sistemas de razonamiento con respecto a las herramientas de visualización, es necesario acompañar este crecimiento con las herramientas acordes. En esta dirección, es importante proveer soporte mediante un lenguaje lógico subyacente mucho más expresivo que redunde en la posibilidad de visualizar diagramas más complejos a partir de inferencias sobre el modelo principal.

Las lógicas para descripción (DLs) son una familia de lenguajes de representación de conocimiento que puede ser usadas para repre-

sentar un dominio de aplicación de una manera estructurada y entendible. A pesar de que su rango de aplicación es amplio (por ejemplo, para razonamiento con esquemas y consultas de Bases de Datos), las DLs son más conocidas como la base para los lenguajes de ontologías tales como DAML+OIL [?] y OWL. Las DLs permiten a los usuarios definir clases o relaciones de los dominios de aplicación en términos de conceptos y roles y, a su vez, restringen la manera en que las clases y relaciones son interpretadas. Asimismo, estas definiciones posibilitan la inferencia de conocimiento implícito a partir de descripciones de conceptos y roles y de la semántica formal de las DLs. Algunos sistemas de razonamiento altamente optimizados, basados en DLs, ya han sido desarrollados, entre los que podemos nombrar a FACT++ [18], RACER [9] y Pellet [15].

Nuestro interés está centrado en enfoques que ofrezcan soporte gráfico y de razonamiento en un framework común. Cualquier representación textual de una ontología es difícil de entender para los usuarios humanos, de manera que su manipulación puede volverse un proceso altamente complejo. La integración de lenguajes gráficos y razonamiento es aún incipiente, por lo tanto, requiere ambientes de desarrollo con esas capacidades. En esta línea de investigación se propone comenzar a completar este *gap* soportando el modelado conceptual gráfico en una manera predictiva.

Finalmente, la motivación principal es satisfacer las necesidades de los expertos del dominio, su entendimiento sobre ontologías y las herramientas de modelado. Los usuarios deberían ver sus modelos junto con sus deducciones expresadas en el mismo lenguaje gráfico.

## 4. Resultados Obtenidos y Trabajo Futuro

Inicialmente, se han analizado diferentes herramientas de desarrollo y se las ha evaluado, teniendo en cuenta fundamentalmente

su interfaz gráfica y el poder expresivo de la lógica subyacente. En este sentido, se tuvo en cuenta la integración de ambas características.

Asimismo, se han estudiado métodos de evolución de ontologías existentes, determinando cómo se identifican los datos para la evolución, a partir de éstos cómo se propagan dichos cambios en la ontología y finalmente, cómo fue implementada la evolución.

A partir de este análisis se ha comenzado con el desarrollo de una metodología que integra el soporte gráfico y el razonamiento formal para la evolución ontológica. Dicha metodología trabaja con la lógica subyacente *SHIQ* y utiliza como núcleo para la evolución consultas del usuario. La metodología ha sido definida formalmente y se mostró su decidibilidad.

Actualmente, nos encontramos trabajando en una extensión de la metodología, formulando nuevas reglas de evolución y verificando su sensatez. Asimismo, se están estudiando la complejidad temporal y espacial.

Como trabajo futuro, proponemos completar la validación de la metodología y, posteriormente, implementarla en una herramienta existente.

## 5. Formación de Recursos Humanos

Uno de los autores de este trabajo está inscripto en el Doctorado en Ciencias de la Computación en la Universidad Nacional del Sur (beca interna doctoral CONICET).

En la Universidad Nacional del Comahue, Facultad de Informática, se otorgaron Becas TICs para finalización de carreras de grado, para alumnos avanzados de la carrera Licenciatura en Ciencias de la Computación. Dos de esas becas, fueron otorgadas a alumnos que desarrollarán sus tesis de grado en esta temática, dirigidos por autores de este trabajo.

Asimismo, se aspira a la formación como investigadores de los miembros más recientes



de ambos grupos.

## Referencias

- [1] T. Berners-Lee, J. Hendler, and O. Lassila. The semantic web. *Scientific American*, May 2001.
- [2] G. Booch, J. Rumbaugh, and I. Jacobson. *Unified Modeling Language User Guide, The (2Nd Edition) (Addison-Wesley Object Technology Series)*. Addison-Wesley Professional, 2005.
- [3] D. Calvanese, G. D. Giacomo, M. Lenzerini, D. Nardi, and R. Rosati. Information integration: Conceptual modeling and reasoning support. In *CoopIS*, pages 280–291. IEEE Computer Society, 1998.
- [4] D. Calvanese, M. Lenzerini, and D. Nardi. Description logics for conceptual data modeling. In *Logics for Databases and Information Systems*, pages 229–263. Kluwer, 1998.
- [5] P. Chen. The entity-relationship model ;toward a unified view of data. *ACM Trans. Database Syst.*, 1(1), Mar. 1976.
- [6] P. Fillottrani, E. Franconi, and S. Tessaris. The new icom ontology editor. In *Description Logics*, CEUR Workshop Proceedings. CEUR-WS.org, 2006.
- [7] P. Fillottrani, E. Franconi, and S. Tessaris. The icom 3.0 intelligent conceptual modelling tool and methodology. *Semantic Web*, 2012.
- [8] G. Guizzardi. *Ontological foundations for structural conceptual models*. PhD thesis, University of Twente, Enschede, The Netherlands, Enschede, October 2005.
- [9] V. Haarslev and R. Möller. Racer system description. In R. Goré, A. Leitsch, and T. Nipkow, editors, *International Joint Conference on Automated Reasoning, IJCAR'2001, June 18-23, Siena, Italy*, pages 701–705. Springer-Verlag, 2001.
- [10] T. Halpin. *Information Modeling and Relational Databases (From Conceptual Analysis to Logical Design)*. Morgan Kaufman, 2001.
- [11] P. Hasse, H. Lewen, R. Studer, and M. Erdmann. The NeOn Ontology Engineering Toolkit. 2008.
- [12] H. Knublauch, R. Ferguson, N. Noy, and M. Musen. The protégé owl plugin: An open development environment for semantic web applications. 2004.
- [13] B. Motik and R. Studer. KAON2—A Scalable Reasoning Tool for the Semantic Web. In *Proceedings of the 2nd ESWC'05*, 2005.
- [14] W. OWL Working Group. *OWL 2 Web Ontology Language: Document Overview*. W3C Recommendation, 27 October 2009. Available at <http://www.w3.org/TR/owl2-overview/>.
- [15] E. Sirin, B. Parsia, B. C. Grau, A. Kalvanpur, and Y. Katz. Pellet: A practical owl-dl reasoner. *Web Semant.*, 5(2):51–53, June 2007.
- [16] P. Spyns, R. Meersman, and M. Jarrar. Data modelling versus ontology engineering. *SIGMOD Rec.*, 31(4):12–17, Dec. 2002.
- [17] TopQuadrant. TopQuadrant — Products — TopBraid Composer, 2011.
- [18] D. Tsarkov and I. Horrocks. Fact++ description logic reasoner: System description. In *In Proc. of the Int. Joint Conf. on Automated Reasoning (IJCAR 2006)*, pages 292–297. Springer, 2006.