

GODO: Generación inteligente de Objetivos para el Descubrimiento de servicios web semánticos*

GODO: Goal driven orchestration for Semantic Web Services

Juan Miguel Gómez

Universidad Carlos III de Madrid
Avda. Universidad 30,
28911 Leganés(Madrid)
juanmiguel.gomez@uc3m.es

Javier Chamizo

Universidad Carlos III de Madrid
Avda. Universidad 30,
28911 Leganés(Madrid)
javier.chamizo@uc3m.es

Resumen: En este artículo se presentará GODO, un motor de última generación que usa el procesamiento de lenguaje natural y técnicas de mapeo para orquestar y alcanzar metas por medio de servicios web.

Palabras clave: web semántica, ontology learning, orquestación

Abstract: In this article we present GODO, a search engine which uses NLP and mapping techniques to orchestrate and achieve goals by means of web services.

Keywords: semantic web, ontology learning, orchestration

1. Introducción

La Web ha cambiado de ser un mero repositorio de información a convertirse en una plataforma de negocio, donde múltiples organizaciones pueden desplegar, compartir y explorar procesos de negocio a través de servicios web. Han surgido así nuevos y prometedores campos de aplicación, como los relacionados con la Web Semántica. En concreto los Servicios Web Semánticos están ayudando a potenciar el despliegue y composición de servicios, pero aún no está completamente resuelto el problema de cómo buscar y descubrir los servicios y hacer un uso sencillo y dinámico de ellos. Además, el hecho de que los usuarios puedan expresar sus deseos en un lenguaje natural, en lugar de tener que conocer un lenguaje lógico complejo, está limitando la eficiencia de la aproximación seguida por los servicios web semánticos.

El presente proyecto pretende conseguir una solución que combinando tecnologías de análisis de lenguaje natural y semántica, consiga que el usuario final de las aplicaciones basadas en servicios web semánticos pueda expresar sus deseos y objetivos de una forma más sencilla al sistema.

2. Objetivo del Proyecto

Como objetivo principal del proyecto GODO, se persigue desarrollar una infraestructura semántica basada en una aplicación de técnicas de análisis de lenguaje natural y ontologías, de manera que pueda ofrecerse al usuario una capa de presentación sencilla para definir sus objetivos de búsqueda y descubrimiento de servicios web semánticos. Actualmente, las plataformas orientadas a los servicios web semánticos demuestran ciertas limitaciones:

A nivel de usuario final, existe la dificultad de definir cuales son los objetivos (goals) para buscar y descubrir los servicios que necesita. Actualmente los servicios web semánticos no ofrecen un interfaz amigable para poder definir los objetivos de búsqueda, y es necesario un gran conocimiento de lenguajes de lógica matemática para poder llegar a definir uno de esos objetivos.

A nivel técnico hay varios problemas:

- Existen varios marcos de definición de servicios web semánticos tales como, WSMO(Cristina Feier, 2005), OWL-S(Martin, 2005) o METEOR-S(Abhijit A. Patil, 2004); cada uno de ellos con su propio lenguaje y definición.
- La definición de goals es a día de hoy un tema complejo que exige un gran conocimiento

* Proyecto GODO(FIT-340000-2007-134) Participantes: Atos Origin S.A., Universidad de Murcia, Universidad Carlos III de Madrid

Aún así, las arquitecturas de *semantic web services* nos facilitan un marco adecuado para la construcción de plataformas de oferta y descubrimiento de servicios. Por otro lado, en los últimos años las plataformas de descubrimiento y ejecución de servicios web semánticos han evolucionado hasta hacerse una realidad, al menos como resultado de proyectos de investigación. Así, proyectos financiados por la Unión Europea, como DIP¹ e INFRAWEBs² han creado una serie de APIs y entornos válidos para crear aplicaciones basadas en servicios web semánticos. Sin embargo, la realidad es que la interacción de usuarios finales y aplicaciones para expresar los objetivos para el descubrimiento de servicios, es aún muy compleja y requiere un gran conocimiento de la plataforma a utilizar y del lenguaje semántico en que se sustenta.

El objeto del proyecto GODO es pues, el de acercar al usuario final a la definición de objetivos para buscar y descubrir servicios web semánticos. En GODO se pretende hacer una realización de este paradigma utilizando una plataforma de servicios web semánticos ya existente, para probar así que la definición de objetivos de una manera sencilla y sin requerir especiales conocimientos de lenguajes complejos de ontologías sea una realidad.

3. Tecnologías PLN y Semánticas

El uso de tecnologías semánticas junto con el Procesamiento del Lenguaje Natural (PLN) permite representar la estructura conceptual del lenguaje, proporcionando una riqueza semántica mayor que un lexicón computacional o un tesoro.

La relación que se establece entre las ontologías y las técnicas de PLN es bidireccional: por un lado son herramientas para la representación de redes semánticas y por otro el PLN es una técnica importante en la construcción automática de ontologías (*ontology learning*).

Además al representar conceptos (no términos), son muy útiles en los sistemas de traducción automática, puesto que a cada concepto se le pueden asociar las formas lingüísticas (una o varias) que lo representen en cada lengua.

Según la Recomendación Estándar Británica para la selección, formación y definición de términos técnicos, *los conceptos*

son constructos mentales, abstracciones que se pueden emplear para clasificar los distintos objetos del mundo en interior. Mientras que los términos son las unidades léxicas concretas que se emplean para referirse a un concepto. No siempre existe un único término para referirse a un concepto, por este motivo ha habido intentos de imponer una relación biunívoca entre término-concepto, atribuyendo un término a cada concepto y representando cada concepto por un solo término. El objetivo es facilitar la comunicación científico-técnica y eliminar la polisemia, sinonimia y homonimia. Esta fue la propuesta realizada por la ISO 704, que por otro lado ha recibido varias críticas.

La aproximación del proyecto GODO se refiere al uso de PLN y las ontologías, centrándose en la construcción automática de ontologías a partir de texto y en la detección de los elementos ontológicos (conceptos, clases, relaciones, atributos) a partir de texto en lenguaje natural.

La construcción automática de ontologías se ha convertido en uno de los principales focos de investigación dentro del ámbito de la Web Semántica. Las ontologías en la Web Semántica se utilizan como estructura compleja para la representación del conocimiento, generalmente de un dominio. Dicho conocimiento está contenido en muchas ocasiones en textos escritos en lenguaje natural. La construcción de ontologías es un proceso lento y costoso que ralentiza el avance de la Web Semántica, por esto es necesario desarrollar métodos eficaces para la generación automática de ontologías a partir de lenguaje natural.

Bibliografía

- Abhijit A. Patil, Swapna A. Oundhakar, Amit P. Sheth Kunal Verma. 2004. Meteor-s web service annotation framework. En *Procs. of the 13th international conference on World Wide Web*.
- Cristina Feier, John Domingue. 2005. Wsmo primer. <http://www.wsmo.org/TR/d3/d3.1/v0.1/>, April.
- Martin, David et al. 2005. Bringing semantics to web services: The owl-s approach. *Lecture Notes in Computer Science*, 3387:26–42.

¹<http://dip.semanticweb.org/>

²<http://www.infrawebs.org/>