

Tecnología semántica, un lenguaje universal para la web

Semantic Knowledge Engineering and Applications fue uno de los cursos de la Escuela de Verano del Departamento de Ingeniería de Sistemas y Computación (DISC) en el 2017. Salvatore Flavio Pileggi (profesor invitado) y María del Pilar Villamil (profesora del DISC), explican las ventajas de la tecnología semántica y advierten que el verdadero reto es aplicarla en el mundo real.

Para un humano es fácil deducir que si María del Pilar es la madre de Alejo, este es el hijo de María del Pilar. Pero para una máquina establecer esa relación en un contexto general como la web no es tan sencillo: son dos afirmaciones separadas que requieren razonar en un contexto de interoperabilidad para llegar a la misma deducción. Es decir, se necesita un estándar capaz de definir unívocamente los conceptos (María del Pilar y Alejo, en este caso) y las relaciones existentes entre ellos (madre-hijo); de conectar los datos, la información y el conocimiento alojados en la torre de Babel que es la web; de definir reglas de inferencia para soportar un razonamiento estándar que pueda otorgar un significado preciso e independiente de interpretaciones individuales.

Al usar tecnología semántica, cada quien podría expresar el mismo concepto de formas distintas —digamos, doctor y médico— y aun así las diferentes partes podrían entenderse, teniendo a la web semántica como una especie de lenguaje universal.

Si dicha tecnología se aplicara en ámbitos que requieren integrar información heterogénea como la salud o los impuestos, una nueva generación de sistemas podría permitir a un médico revisar la historia clínica completa de un paciente con datos de diferentes fuentes; el Ministerio

de Salud tomaría decisiones basadas en el todo y no en reportes individuales, o la Dian conocería los bienes y los movimientos financieros de un ciudadano y los cruzaría con otras oficinas de impuestos en el exterior.

Con esa explicación los profesores Salvatore Flavio Pileggi y María del Pilar Villamil introducen el tema del conocimiento semántico que abordaron en el curso *Semantic Knowledge Engineering and Applications*, organizado por el DISC, dentro de la Escuela de Verano 2017. Este se desarrolló entre el 27 de junio y el 14 de julio y su propósito era ir más allá de lo meramente técnico, proporcionando las herramientas y una metodología basada en la conceptualización que soporta el proceso de Ingeniería del Conocimiento, desde cero hasta su implementación en entornos complejos.

Ambos explican que la tecnología semántica va más allá de estructurar los datos. Al contrario de la web social (o web 2.0) que proporciona soportes tangibles como la posibilidad de acceder a contenidos multimediales o interactuar a través de redes sociales a gran escala, la tecnología semántica, a través de metadatos, actúa de forma totalmente imperceptible para el usuario final en múltiples aplicaciones que usamos a diario (motores de búsqueda por ejemplo).



Foto: Aldair Morales, DISC.

Salvatore Flavio Pileggi es Ph.D. en Comunicaciones de la Universidad Politécnica de Valencia e investigador de la School of Information Technology and Electrical Engineering (ITEE) en University of Queensland (Brisbane, Australia). **María del Pilar Villamil** es doctora en Informática del Institut National Polytechnique de Grenoble (Francia) y profesora del DISC.

Los retos

Cuando comenzó el curso de Escuela de Verano, Salvatore Flavio Pileggi y María del Pilar Villamil les preguntaron a los participantes qué sabían del tema. Eran ingenieros que trabajan en empresas colombianas, muchos de ellos con maestría, y lo desconocían por completo o casi.

Ese resultado reafirma el reto de lograr que los ingenieros actuales y los que están formándose vean el potencial de la tecnología semántica y la utilicen en su día a día para que siga madurando.

El primer paso es construir ese significado de los datos y el siguiente será cómo descubrir patrones sobre esa semántica. Es decir, una vez se tengan los significados y se logre la interacción, se debe poner una capa encima y analizarla para hallar elementos comunes o detectar qué falta.

Imágenes: www.pixabay.com

¿Tecnología en consolidación o con impacto limitado?

Los profesores Pileggi y Villamil destacan que la web semántica se puede considerar un entorno relativamente maduro, teóricamente listo para su aplicación en el mundo real e incluso ya hay evidencias claras de que se está evaluando la adopción de modelos semánticos complejos (ontologías) en casos reales. Así sucede, por ejemplo, con muchos gobiernos interesados en datos abiertos, en facilitar la integración y la posibilidad de compartir información entre las agencias y en el análisis de datos heterogéneos. O también estudios recientes demuestran una presencia masiva de metadatos semánticos en páginas web comunes, aunque el uso de la tecnología semántica es bastante básico y no incluye técnicas de razonamiento avanzadas.

No sucede igual con las empresas, especialmente las multinacionales, que parecen poco entusiasmadas en adoptarla. Según el profesor Pileggi, aunque es difícil proporcionar un análisis exhaustivo, todo indicaría que muchas asocian el uso de una tecnología abierta con un peligro para

su negocio y no evidencian en ella nuevas oportunidades.

A ese desinterés contribuye el hecho de que a menudo los investigadores se mueven en un círculo: plantean problemas que no siempre están alineados con los de las organizaciones y luego los resuelven; esto dificulta la transferencia de soluciones al sector productivo.

Por lo anterior, el profesor Flavio Pileggi señala que la tecnología semántica parece enfrentarse a un momento crucial, pues otras podrían superarla e incluso necesita nuevas energías en términos de investigación porque tiene limitaciones. Y concluye: “Los datos abiertos y la integración de datos heterogéneos son muy populares entre los gobiernos y por eso la investigación sobre semántica se mantiene con vida; pero sin un esfuerzo para transferir soluciones al sector productivo, la tecnología semántica, en lugar de implementar la tercera versión de la web con impacto masivo y global, podría reducirse a un paréntesis, a algo que ha sido muy popular en términos de investigación durante un período relativamente largo y luego ha reconsiderado drásticamente sus aspiraciones reales”. ■

“A menudo los investigadores se mueven en un círculo: plantean problemas que no siempre están alineados con los problemas de las organizaciones y luego los resuelven; esto dificulta la transferencia de soluciones al sector productivo”.
Salvatore Flavio Pileggi