

LENGUAJE DE ONTOLOGÍAS WEB APLICADAS AL SECTOR TURÍSTICO

Alexis Messino Soza y Luis C. Tovar Garrido
Universidad Simón Bolívar, Barranquilla (Colombia)

Resumen

El presente artículo describe el proyecto de investigación titulado “Plataforma TIC para propiciar la organización y colaboración de los diferentes actores del sector turístico en el Caribe Colombiano” y desarrolla algunos conceptos sobre aplicaciones de Ontologías Web al sector turismo. Una ontología define los términos a utilizar para describir y representar un área de conocimiento. Las ontologías son utilizadas por personas, bases de datos y aplicaciones que necesitan compartir un dominio de información común e incluyen definiciones de conceptos básicos del dominio y las relaciones entre ellos.

El artículo inicia con el análisis del concepto de ontologías, los componentes que lo conforman y algunos lenguajes de ontologías web frecuentemente utilizados. La segunda parte explica cómo se pueden aplicar las ontologías al sector turístico. El artículo finaliza explicando cómo se construyó un buscador de atractivos turísticos basado en ontologías.

Palabras claves: Ontología, OWL, turismo, buscador, conocimiento.

Abstract

This paper describes the research project named “ICT Platform to promote organization and cooperation between several actors in the Colombian Caribbean tourism sector”; it develops some concepts on Web Ontology applied to the tourism sector. Ontology defines the terms used to describe and represent an area of knowledge. Ontology is used by people, databases, and applications that need to share a common information domain, and it includes definitions relevant to basic domain concepts, and their relationship.

The paper starts with the conceptual analysis of ontology, its structural components, and some commonly used web ontology languages. The second part explains how to apply it to the tourism sector, and finishes by explaining how the touristic attractions search engine based on ontologies was built.

Keywords: Ontology, OWL, tourism, searcher, knowledge.

Introducción

Al hablar de ontologías es inevitable pensar en la web semántica, la cual se basa en la idea de añadir metadatos semánticos y ontológicos a la *world wide web*. Esa información adicional -que describe el contenido, el significado y la relación entre los datos- se debe proporcionar de manera formal a los agentes humanos o autómatas encargados de su procesamiento.

Actualmente, el Internet es una tecnología de uso cotidiano por todos los actores en una sociedad, comparable en cobertura a otros medios de comunicación tradicionales como la radio, la televisión o el teléfono, a los que supera en numerosos aspectos. La web es hoy un medio extraordinariamente flexible y económico para la comunicación, el comercio y los negocios, así como para el entretenimiento, el acceso a información y la oferta de servicios, entre otros aspectos. El ritmo vertiginoso de crecimiento de la web, es causante de que las tecnologías sobre las que se basa, hayan experimentado una rápida evolución.

Algunas de esas tecnologías que hacen posible la web semántica incluyen lenguajes para la representación de ontologías, *parsers*, lenguajes de consulta, entornos de desarrollo, módulos de gestión (almacenamiento, acceso, actualización) de ontologías, módulos de visualización, conversión de ontologías, y otras herramientas y librerías.

El artículo está organizado de la siguiente forma: inicialmente se definen algunos de los conceptos básicos empleados. Seguidamente, se ofrece una descripción del uso de OWL® (Web Ontology Language) en el sector turístico. Posteriormente, se explora la implementación de un motor de búsqueda basado en ontologías y finalmente, se presentan las conclusiones.

Conceptos básicos

Ontología

En el contexto de información, una definición del término “ontología” ampliamente adoptada es la siguiente: *Una ontología es un compromiso explícito y formal de especificación de una conceptualización compartida.* (Thomas, 1995).

La principal función de una ontología es representar de manera formal conceptos atribuibles a la solicitud de dominio, es decir, la porción de la realidad en estudio, el conocimiento existente. Especificar la conceptualización significa asignar un significado inequívoco a fin de que se definan los conocimientos en un dominio específico. La palabra compartida refleja el hecho de que la ontología capture el conocimiento consensual, o aceptado por un grupo.

Una ontología no es un conocimiento del mundo objetivo, sino que representa el punto de vista compartido por un grupo (Hendler, 2004) (Guizzardi, 2007). Tiene por objeto eliminar, o al menos reducir, la confusión conceptual y de la terminología, con el fin de tener una interpretación compartida (Hendler, 2004), (Meersman *et al.*, 2002).

Normalmente una ontología se compone de:

- Conceptos que describen un dominio de aplicación específico, como el turismo, el tema de este documento.
- Definiciones de los conceptos, a través de un lenguaje formal o lenguaje natural.
- Las relaciones entre los conceptos expresados a través de un lenguaje formal. Las relaciones pueden ser de varios tipos: taxonomía (es un), meronimia (es parte de), asociaciones, etc.
- Conceptos generados que expresan las principales categorías en las que se organiza el mundo, lo que constituye el más alto nivel posible de una ontología del mundo. Sucede en general, independiente del dominio.
- Una ontología puede ocurrir en diversas formas y tener diferentes niveles de formalización, pero siempre incluye un vocabulario de términos con la descripción de su significado (Fernández *et al.*, 2003).

Componentes de la ontología

Las ontologías tienen los siguientes componentes que sirven para representar el conocimiento de algún dominio (Honrubia, 2002):

- **Conceptos:** son las ideas básicas que se intentan formalizar. Los conceptos pueden ser clases de objetos, métodos, planes, estrategias, procesos de razonamiento, etc.

- **Relaciones:** representan la interacción y enlace entre los conceptos del dominio. Suelen formar la taxonomía del dominio.
- **Funciones:** son un tipo concreto de relación donde se identifica un elemento mediante el cálculo de una función que considera varios elementos de la ontología.
- **Instancias:** se utilizan para representar objetos determinados de un concepto.
- **Axiomas:** son teoremas que se declaran sobre relaciones que deben cumplir los elementos de la ontología. Por ejemplo: “Si A y B son de la clase C, entonces A no es subclase de B”, “Para todo A que cumpla la condición CI, A es B”, etc.

Estos últimos componentes -los axiomas- permiten, junto con la herencia de conceptos, inferir conocimiento que no esté indicado explícitamente en la taxonomía de conceptos.

Gráfica 1. Ontología en el lenguaje OWL DL

```
<?xml version="1.0"?>
<rdf:RDF
  xmlns:lugaresGeograficos="http://www.owl-ontologies.com/Ontology1218023717.owl#"
  xmlns:transportes="http://www.owl-ontologies.com/Ontology1218023205.owl#"
  xmlns:swrl="http://www.w3.org/2003/11/swrl#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:proveedores="http://www.owl-ontologies.com/Ontology1218019535.owl#"
  xmlns:xsp="http://www.owl-ontologies.com/2005/08/07/xsp.owl#"
  xmlns:swrlb="http://www.w3.org/2003/11/swrlb#"
  xmlns:precios="http://www.owl-ontologies.com/Ontology1218024004.owl#"
  xmlns:protege="http://protege.stanford.edu/plugins/owl/protege#"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
  xmlns="http://www.owl-ontologies.com/Ontology1218125991.owl#"
  xml:base="http://www.owl-ontologies.com/Ontology1218125991.owl">
  <owl:ontology rdf:about="">
    <owl:imports rdf:resource="http://www.owl-ontologies.com/Ontology1218023717.owl"/>
    </owl:ontology>
    <owl:Class rdf:ID="Republicanos">
      <rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
        >Representa todos los sitios arquitectonicos que son republicanos.</rdfs:comment>
      <rdfs:subClassof>
        <owl:Class rdf:ID="SitiosArquitectonicos"/>
      </rdfs:subClassof>
    </owl:Class>
    <owl:Class rdf:ID="CentrosTecnicosCientificos">
      <rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
        >Representa todos los sitios que son tecnicos-cientificos</rdfs:comment>
      <rdfs:subClassof>
        <owl:Restriction>
          <owl:onProperty>
            <owl:DatatypeProperty rdf:ID="horario"/>
          </owl:onProperty>
          <owl:minCardinality rdf:datatype="http://www.w3.org/2001/XMLSchema#int">
            >1</owl:minCardinality>
          </owl:Restriction>
        </rdfs:subClassof>
        <rdfs:subClassof>
          <owl:Class rdf:ID="SitiosInteres"/>
        </rdfs:subClassof>
      </owl:Class>
```

Lenguaje de ontología web

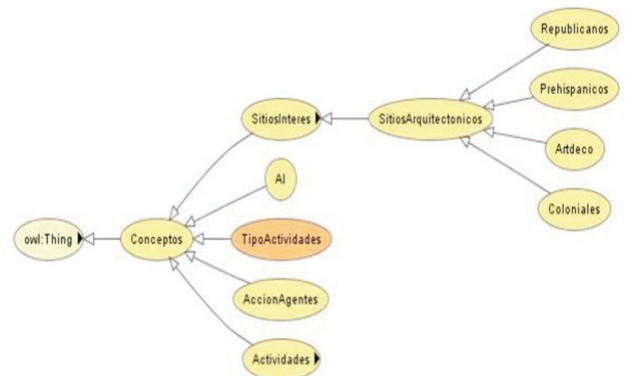
Las ontologías se pueden desarrollar en diferentes lenguajes como RDF®, RDF SCHEMA®, XML®, XML SCHEMA® y OWL®. Este último es el más utilizado por tener mayor capacidad para expresar significado y semántica; OWL® va más allá de los otros lenguajes mencionados, en su capacidad para representar contenido interpretable por un computador en la web (Hendler, 2004).

Actualmente, OWL® posee tres versiones OWL Lite®, OWL DL® y OWL Full®; estas versiones incorporan diferentes funcionalidades. *OWL Lite®* está construido de tal forma que toda sentencia pueda ser resuelta en tiempo finito; la versión más completa de *OWL DL®* puede contener ‘bucles’ infinitos y se basa en la lógica descriptiva, mientras que *OWL Lite®* se basa en una lógica menos expresiva (Hendler, 2004). Un ejemplo del código generado en OWL® se puede apreciar en la gráfica 1.

Ontologías aplicadas al sector turístico

Existen investigaciones sobre la manera como se describen los conceptos relacionados con el ámbito del turismo en una ontología. Es evidente que este ámbito es amplio y los límites no están bien definidos. El propósito de una ontología turística es facilitar la descripción de los atractivos turísticos que se ofrecen como un servicio web; por lo tanto, los conceptos que se han elegido para tener en cuenta son principalmente los buscadores de sitios de interés, hoteles, operadores turísticos y actividades que se puedan realizar en determinado lugar (Batzold *et al.*, 2004), como se puede apreciar en el diagrama de la gráfica 2.

Gráfica 2. Conceptos de una Ontología Turística



En lo que se refiere al hallazgo de contenidos para identificar conceptos de ontología en ámbito del turismo, se ha estudiado todo “el ciclo de vida” de un viaje considerando todas las necesidades de un turista. Por esta razón, se analizaron los servicios de reserva de vehículos y las instalaciones en las cuales pueden alojarse los turistas; también los servicios complementarios, como predicción meteorológica, cálculo de rutas, información sobre los países y las ciudades, entre otros.

Algunas dificultades encontradas se relacionan con determinados conceptos de turismo que no están bien definidos, por ejemplo, con respecto a los alojamientos, en algunos países no existe una legislación nacional, sino normas regionales.

Para resolver estas dificultades se consultaron la mayor cantidad de fuentes autorizadas identificables y se extrajeron los conceptos que les eran comunes (Pandolfo *et al.*, 2004).

Durante la identificación y el análisis de los conceptos relacionados con en el turismo como dominio, se llegó a la conclusión que no es posible aislar estos conceptos de otros que tienen que ver con el llamado “sentido común” y que son independientes del dominio particular que se trate. Como ejemplo podemos citar: los relacionados con el tiempo (día y fecha), la ubicación del espacio (lugar, dirección, etc.), entre otros.

Esa familiaridad puede ser diseñada y relacionada de una manera diferente. Si se analizan los distintos significados de los términos que se utilizan en el idioma inglés para describirlos, se descubre que un término que parece detallar intuitivamente un solo sentido en la designación de diversos conceptos es diferente el uno del otro. Es entonces cuando se hace necesario tomar decisiones, tratando de identificar entre las diferentes interpretaciones, las que son más pertinentes para el ámbito del turismo (Pandolfo *et al.*, 2004).

El significado de los conceptos se expresa principalmente a través de las definiciones en lenguaje natural. Con este fin se puede utilizar OWL DL®. A través de este lenguaje se expresan las relaciones existentes entre los diferentes conceptos. El uso de una ontolo-

gía superior, sobre todo cuando no tiene una amplia utilización de herencia múltiple, tiene la ventaja de facilitar una navegación manual de la ontología y, por tanto, simplificar la recuperación de conceptos que pueden servir para una descripción. La construcción de una ontología es una tarea aún más delicada que la elección de los conceptos de dominio (Batzold *et al.*, 2004), (Pandolfo *et al.*, 2004).

Construcción de un buscador de atractivos turísticos basado en ontologías

La forma tradicional de encontrar resultados en motores de búsqueda robustos es por medio de comparaciones de texto realizadas entre las páginas html de los sitios previamente almacenados e indexados en su base de datos (si el sitio web no es de mucha importancia, solo se guardan de una a dos páginas); procedimientos de motores de búsqueda menos complejos usan la información puntual de los sitios (títulos, palabras claves) como criterio de resultado o la concurrencia sobre un dominio, es decir, las veces que se accede a ella (implementado por Google). Esta forma de manejar las búsquedas sin ningún tipo de categorización, sino solo con un texto plano como referencia, crea resultados inexactos en muchos casos (Cruz *et al.*, 2002).

Las búsquedas ontológicas, basadas en esquemas conceptuales, generan resultados coherentes con el criterio del usuario sobre uno o varios dominios de información. El dominio es el conjunto de conceptos que se definen por medio de la ontología. A continuación se presentan los pasos que se siguieron para la implementación de un dominio de información creado bajo el lenguaje de ontologías OWL sobre los *atractivos turístico del Caribe Colombiano* en el buscador web del Portal “Somos Caribe colombiano... mágica realidad” (www.somoscaribe.com) y los resultados obtenidos a partir de esta experiencia.

Comprender el dominio del concepto de la información

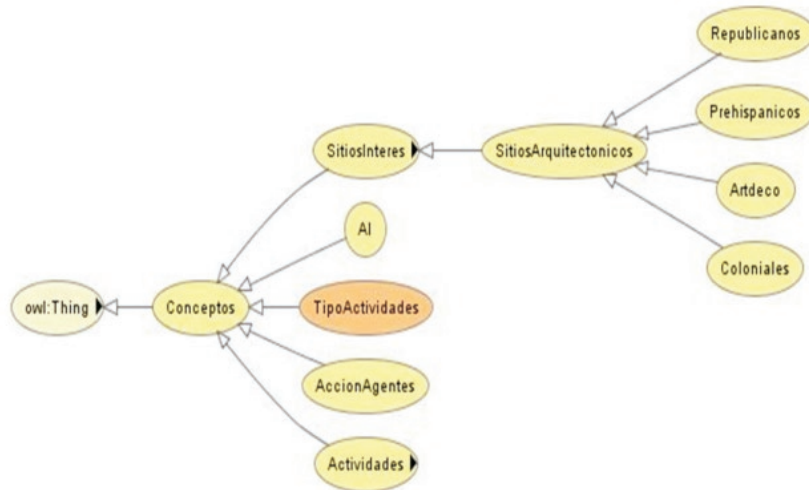
La comprensión del dominio del concepto de la información es lo más importante en la creación

de una ontología web; en este caso al tener la base de conocimiento (ontología e instancias de las clases) ya creada, se debe comprender el esquema conceptual que se va a utilizar. Los editores ontológicos permiten comprender el dominio del concepto de una forma fácil y clara. Existen diferentes editores (swoop®, semantic turkey®, Sinaptyc®, entre otros.), en el caso del buscador del portal turístico www.somoscaribe.com se utilizó el editor PROTÉGÉ® creado por la Universidad de

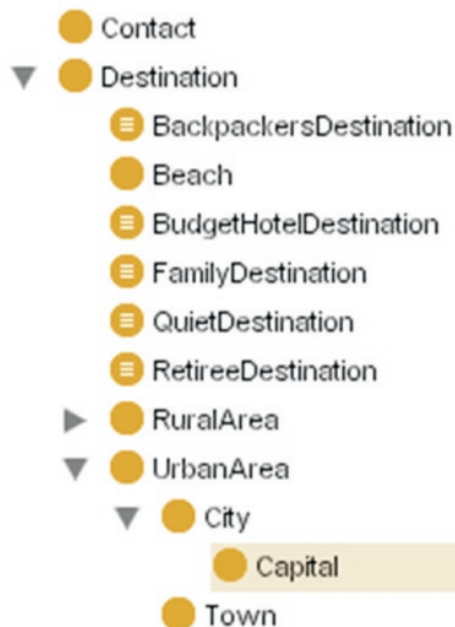
Stanford (<http://protege.stanford.edu/>), que provee gran soporte y además está desarrollado sobre plataforma libre.

PROTÉGÉ® entrega diferentes formas de visualizar los conceptos y relaciones. Entre ellas tenemos: Árbol jerárquico de las clases y subclases de la ontología (Gráfica 4), plugins como OWLViz® y Jambalaya® que muestran gráficamente nuestro dominio (Gráfica 5, Gráfica 3).

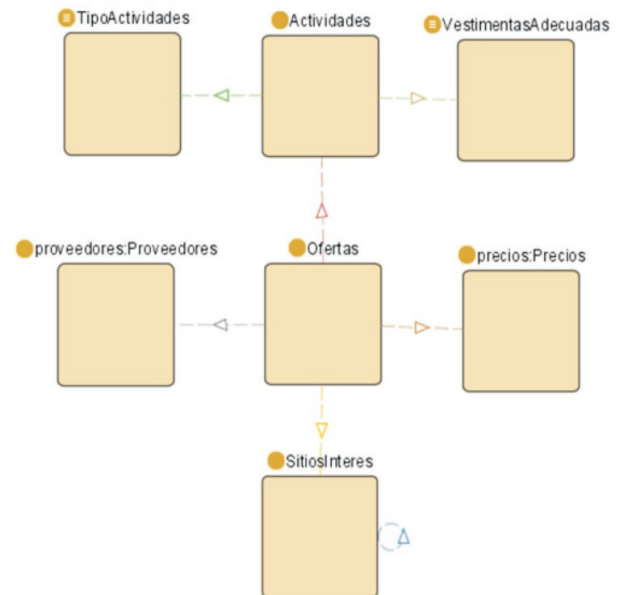
Gráfica 3. Ontología creada en PROTÉGÉ 3.4



Gráfica 4. Árbol jerárquico, editor ontológico PROTÉGÉ®



Gráfica 5. Gráficos del dominio *atractivos turísticos* con Jambalaya® y OWLViz®



Una vez entendido bien el dominio, se deben limitar los conceptos importantes para alcanzar el objetivo deseado.

Definir conceptos relevantes de las búsquedas

Definir de forma adecuada cuáles serán los criterios de búsqueda es fundamental en la creación del buscador ontológico. El colocar rastreo dentro de propiedades de datos que no sean relevantes o poco precisas puede crear resultados inconsistentes. Primero que todo, se debe tener claro el fin o propósito que se le dará a la ontología. En el caso particular del buscador turístico, su propósito es mostrar lugares de interés relacionándolos con el nombre, municipio o departamento donde están ubicados y con las características del atractivo. Esto permite definir criterios con relevancia para la búsqueda:

1. Nombre de atractivo turístico.
2. Atractivos turísticos que se encuentren dentro de un municipio (nombre de municipio).
3. Atractivos turísticos que se encuentren dentro de un departamento (nombre del departamento).
4. Sinónimos del atractivo turístico (utilizado en casos como Barranquilla y cuyo sinónimo es Puerta de Oro).
5. Información dentro de una descripción corta del atractivo turístico.

Estos son los 5 criterios de búsqueda manejados en la búsqueda de atractivos y que van clasificados según un rango de importancia. Dentro de la ontología de atractivos turísticos se encuentran otras propiedades de la instancia (descripciones ampliadas, fecha y hora de atención) que no son importantes o pueden crear resultados menos exactos.

Utilización de herramientas para el manejo de ontológicas

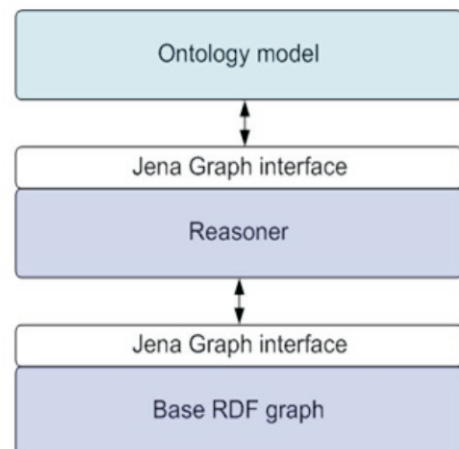
El manejo de la ontología se debe hacer por medio de una herramienta que permita crear el esquema conceptual en memoria, hacer razonamientos, consulta de datos, modificar y crear nuevas relaciones o conceptos, entre otros procesos. Un Framework apropiado es JENA® de la plataforma Java, que en su API (interfaz de programación de aplicaciones) contiene diferentes tipos de razonadores y varios

tipos de variables para el manejo de ontologías, tales como los `ontmodel`, que contienen la ontología cargada en memoria, o los `ontProperties` para obtener propiedades. A su vez, permite el uso de lenguajes de consulta dentro de la web semántica (SPARQL®), como se aprecia en la gráfica 6.

Generación de consultas por medio del lenguaje SPARQL

Para el manejo y consulta de datos dentro de una ontología se usó SPARQL® (*SPARQL Protocol and RDF Query Language*). El bloque básico de construcción de una consulta SPARQL® es el patrón de tripleta. Una tripleta puede tener variables en lugar de fuentes en cualquiera de estas tres posiciones, sujeto, predicado y objeto. Las variables son indicadas como símbolos precedidos por un caracter especial: ? (Allemnag *et al.*, 2008).

Gráfica 6. Ontología y razonadores en JENA



Un ejemplo sobre la ontología de atractivos turísticos es:

```

PREFIX st: <http://purl.org/dc/elements/1.1/>
SELECT ?nombres
WHERE {
  ?sitios st:nombre ?nombres
}
  
```

Se puede observar la sintaxis utilizada en SPARQL®. El PREFIX (prefijos) se utiliza para almacenar una URL con la que se va a referenciar una ontología, en este caso será sobre los atractivos turísticos y tendrá el

prefijo st. SELECT son los datos que serán devueltos por la búsqueda; en nuestro caso todos los nombres de los atractivos o sitios turísticos. WHERE son las condiciones que debe cumplir el dato para poder estar dentro de los resultados.

En general, se puede definir la consulta como: qué objetos (?sitios) que estén dentro de la ontología de atractivos turísticos tienen un nombre (dc:nombre) cualquiera (se dice que cualquiera porque la propiedad está precedida por una variable ?nombres). El resultado podría ser:

Nombres
“El Rodadero”
“Parque Tayrona”
“El Hoyo soplador”

Los resultados obtenidos por SPARQL® son solo de tipo literal, es decir solo los datos que se han insertado a las instancias de las clases.

Conclusión

Se logra observar que el sector turístico es un campo en el cual se pueden aplicar ontologías web. El empleo de ésta es adecuado para el desarrollo de servicios automatizados a usuarios como puede ser el caso concreto de servicios turísticos.

Existen actualmente diversos sistemas y prototipos que ofrecen servicios turísticos como por ejemplo restaurantes o museos. En estos casos se suelen tratar los servicios de una forma independiente y no integrada, como sería si se utilizara ontologías web.

La búsqueda basada en ontologías es un paso importante en la generación de resultados más precisos y con niveles mayores de exactitud en la información; esto es aplicable a dominios turísticos. El uso e inversión en el desarrollo de lenguajes de ontologías promete un futuro esperanzador (Elmundo.es, 2008).

Cabe destacar, que en la web se pueden encontrar ontologías desarrolladas y mejoradas por diferentes personas, que pueden ser usadas con múltiples propósitos; la propuesta es organizar un único dominio que sea universal y a su vez contenga otros dominios entre sí, para poder crear un esquema completo de toda la información en la web.

Las principales ventajas de esta nueva revolución en Internet serán el desarrollo de aplicaciones con esquemas de datos comunes y búsqueda de información con inferencias. Para poder lograr estos propósitos se necesita unificar los contenidos semánticos por medio de ontologías que formalicen este conocimiento de forma consensuada y reutilizable. Se necesita un lenguaje común basado en web, con suficiente capacidad expresiva y de razonamiento para representar la semántica de las ontologías.

Referencias

- Allemlag, D. and Hendler, J. (2008). Semantic web for the working Ontologist. Estados Unidos, pp. 85-87.
- Batzold, M.; Navarro, J.; Botti, V. et al. (2004). Desarrollo de servicios turísticos a usuarios. Departament de Sistemes Informatics i Computacio. Universitat Politècnica de Valencia, Spain.
- Cruz, I.; Decker, S.; Euzerat, J.; McGuinness, D. (2002). The emerging semantic web. Ciudad: Estados Unidos, pp. 154-155.
- Elmundo.es (2008). Bruselas quiere que los Veintisiete lideren el desarrollo de la Web 3.0. Consultado el 15 de octubre de 2009 en <http://www.elmundo.es/navegante/2008/09/30/tecnologia/1222771475.html>
- Fernández, M; y Gómez, A. (2003). The integration of OntoClean in WebODE. Facultad de Informática. Universidad Politécnica de Madrid, pp. 2-8.
- Guizzardi, G. (2007). On Ontology, ontologies, Conceptualizations, Modeling Languages, and (Meta) Models. Federal University of Espírito Santo (UFES), Vitória, Brazil. Laboratory for Applied Ontology (ISTC-CNR), Trento, Italy.
- Hendler, J. (2004). Frequently Asked Questions on W3C's Web Ontology Language (OWL).
- Honrubia, F. (2002). Introducción a las Ontologías. Escuela Universitaria Politécnica de Albacete.
- Lladó, E; Salamanca, I. (2004). Redes Peer-To-Peer para la Gestión de Conocimiento en Destinos Turísticos.

- Fundación IBIT. Palma de Mallorca (Baleares, España).
- Meersman, R.; Spyns, P.; y Jarrar, M. (2002). Data modeling versus Ontology engineering. SPECIAL ISSUE: Special section on semantic web and data management. Vol. 31, pp. 12-17
- Pandolfo, C; Torelli, F. (2004). Ontología del dominio turístico. Fondo investimenti per la Ricerca di Base (FIRB). Programma Strategico Tecnologie abilitanti per la Società della Conoscenza –ICT.
- Thomas, G. (1995). Toward Principles for the Design of Ontologies Used for Knowledge Sharing. International Journal Human-Computer Studies Vol. 43, Issues 5-6, pp. 907-928.

Sobre los autores

Alexis Messino Soza.

Ingeniero de Sistemas. Especialista en Alta Gerencia. Doctorando en Ingeniería de Software, Docente Investigador de la Universidad Simón Bolívar y líder del grupo de investigación IngeBioCaribe. Barranquilla, Colombia.
amessino@unisimonbolivar.edu.co

Luis Carlos Tovar Garrido.

Ingeniero de Sistemas. Especialista en Gerencia de Sistemas de información. Magister en ciencias computacionales. Doctorando en Ingeniería de Software, Miembro del grupo de investigación IngeBioCaribe de la Universidad Simón Bolívar. Barranquilla, Colombia.
ltovar@unisimonbolivar.edu.co

Los puntos de vista expresados en este artículo no reflejan necesariamente la opinión de la Asociación Colombiana de Facultades de Ingeniería.