

# Las Ontologías y el Lenguaje OWL 2

**David E. Narváez**

Universidad Tecnológica de Panamá, estudiante  
david.narvaez@computer.org

**Resumen-** *En la era tecnológica actual, el manejo estructurado de la información toma un papel de suma relevancia debido a la gran cantidad de datos accesibles al público: hemos llegado a tal punto que el procesamiento de toda la información disponible ya no es una tarea alcanzable para el ser humano.*

*De esta realidad nace la necesidad de proveer a las computadoras de la capacidad de procesar información arbitraria, sin embargo, para estos fines, es también necesario estructurar la información de una manera adecuada para ser procesada por las computadoras. Las Ontologías buscan precisamente proveer a las Ciencias de la Información de una estructura común para esta tarea.*

**Palabras Claves-** Ciencias de la Información, Ciencias Computacionales, Ontologías, OWL, Complejidad Algorítmica

## 1. Introducción

Tal vez la distinción más importante que se debe hacer al tratar con cualquier tema de las Ciencias de la Información es aquella que existe entre los términos datos, información y conocimiento. Los datos proporcionan valores como números y textos, que en conjunto componen la información. Sin embargo, la manera en la que la información debe ser interpretada y la manera en que se debe reaccionar ante cierta información es lo que se denomina conocimiento. De la misma manera en que contamos con los tipos de datos para describir los datos, y contamos con las bases de datos para describir la información, las ontologías son las que nos permiten modelar de manera estructurada el conocimiento.

El término ontología fue utilizado originalmente por los pensadores griegos, y se refería al estudio de los objetos que conforman toda la realidad, la categorización de los mismos y la relación que existe entre ellos. En concordancia con la definición original, las ontologías en las Ciencias de la Información definen relaciones y jerarquías entre los objetos que forman el conocimiento y define un vocabulario relacionado con el mismo, como lo muestra el ejemplo que se desarrolla detalladamente en [1].

Una ontología así definida es un objeto abstracto, sin embargo, en la práctica, existen una gran cantidad de lenguajes especialmente diseñados para describir una ontología en un documento digital, facilitando así el procesamiento automatizado de la información. En este artículo trataremos sobre el Web Ontology Language 2 (OWL 2), su historia, sus diferentes niveles y aplicaciones y algunas librerías de software que nos permiten trabajar con ontologías definidas en este lenguaje.

## 2. Historia

Sin duda alguna, el reciente interés por las Ciencias de la Información dentro de las Ciencias Computacionales se debe en gran parte a la popularización del Internet y toda la información

en él disponible. Sin embargo, los esfuerzos por definir un formato común para el intercambio de conocimientos por medio de computadoras datan de mucho antes. A continuación presentaremos una breve reseña de la historia del lenguaje OWL el cual nos permitirá poner en contexto las ventajas y desventajas del mismo.

### 2.1 El Esquema RDF

Con el fin de proveer a las Ciencias Computacionales de un estándar para la representación de ontologías, el World Wide Web Consortium (W3C) definió en 1999 el Marco de Descripción de Recursos (RDF, por sus siglas en inglés) el cual consiste en un conjunto de especificaciones relacionadas con la definición digital del conocimiento [2].

En particular, este marco contiene el Esquema RDF, (RDFS) el cual consiste en un lenguaje para definir conceptos y relaciones entre conceptos, como jerarquías y pertenencias [3].

Este lenguaje pertenece a la familia de lenguajes de marcado, familia a la cual pertenecen también otros lenguajes populares como HTML y XML, este último estrechamente relacionado con RDFS.

El conocimiento en RDF es representado mediante ternas RDF, las cuales son expresiones de la forma Sujeto – Predicado – Objeto en donde el Sujeto y el Objeto son conceptos de la ontología y el Predicado define una relación entre ellos.

### 2.2 OWL

A pesar de que, desde su publicación, RDFS fue ampliamente utilizado para la representación del conocimiento, el mismo posee graves limitaciones como la incapacidad de definir conjunciones de clases y la incapacidad de definir relaciones de cardinalidad. Con el fin de atender estos problemas, la W3C conformó un grupo de trabajo que, basándose en un lenguaje alternativo (también de marcado) para la definición de conocimientos llamado DAML+OIL, publicó OWL, una familia de lenguajes que extendía la funcionalidad presente en el RDFS, la cual fue aceptada como una Recomendación de la W3C el 10 de Febrero de 2004 [4].

Una de las características más importantes de esta especificación es la definición de tres sublenguajes dentro de la misma: OWL Lite, OWL DL y OWL Full, aquí listados en orden creciente de complejidad, que permiten al usuario dotarse de un subconjunto de reglas sintácticas adaptado a la medida de sus necesidades: OWL Lite proporciona un lenguaje sencillo para quienes solamente necesitan categorizar objetos y definir relaciones entre ellos; OWL DL extiende OWL Lite para permitir mayor expresividad, garantizando a la vez que ciertas propiedades lógicas se mantengan dentro de la ontología definida y finalmente OWL Full proporciona todos los grados de libertad posibles en la definición de las ontologías sin garantizar que se mantengan las propiedades lógicas deseables para el razonamiento automatizado sobre la ontología.

### 2.3 OWL 2

Con el fin de mejorar la especificación original del lenguaje OWL, un grupo de investigadores, principalmente de la Universidad de Manchester, propuso el 19 de Diciembre de 2006 una extensión del mismo llamada OWL 1.1

Esta extensión tenía como finalidad la creación de un nuevo grupo de trabajo para adicionar nuevas características como mayor expresividad sintáctica, extensiones sobre los tipos de datos y capacidad de anotaciones, metamodelado simple, entre otras. La conformación de este grupo de trabajo produjo la publicación y aceptación de la especificación OWL 2 como nueva Recomendación W3C para lenguajes de descripción del conocimiento [5].

### 3. ¿Qué es OWL 2?

A continuación detallamos en qué consiste la especificación del lenguaje OWL 2. Veremos también cuáles son los distintos sublenguajes que conforman esta especificación y los componentes que define la misma.

#### 3.1 Definición Formal

Como ya mencionamos, OWL 2 es un lenguaje para la descripción del conocimiento, con un enfoque específico en definir ontologías para la Web. Más específicamente, el objetivo principal del lenguaje es el permitir el procesamiento de la información disponible en la Web por medio de computadoras. Un documento OWL 2 es normalmente compartido como un documento RDF/XML, sin embargo existen otras sintaxis y herramientas de serialización que pueden ser utilizados para representar una ontología descrita en OWL 2 en un documento, a saber, el serializador Turtle, la especificación OWL 2 XML y la Sintaxis Manchester.

La especificación de OWL 2 define de manera abstracta la estructura de una ontología descrita por medio de este lenguaje, sin embargo, existen dos definiciones semánticas distintas que indican cómo interpretar una ontología definida en OWL 2, la Semántica Directa (conocida informalmente como OWL 2 DL o Sintaxis Funcional) y la Semántica Basada en RDF (conocida informalmente como OWL Full).

#### 3.2 Sublenguajes de OWL 2

Al igual que su predecesor, OWL 2 define sublenguajes, también conocidos como Perfiles, los cuales se distinguen entre sí por las propiedades computacionales y lógicas que los mismos garantizan y son subconjuntos del lenguaje OWL DL [6]. Estos perfiles son:

- OWL EL: Este perfil garantiza que todas las tareas de razonamiento comunes dentro de la ontología requieren un tiempo polinomial. Esto es útil, particularmente, al trabajar con ontologías de gran tamaño, pues bajo estas circunstancias es posible intercambiar la expresividad del lenguaje por una mayor eficiencia en los algoritmos.
- OWL QL: Este perfil permite la realización de consultas conjuntivas a la ontología por medio de lenguajes relacionales como SQL garantizando que las mismas puedan ser resueltas en Espacio Logarítmico. Puede ser utilizado para definir ontologías que luego serán consultadas directamente a modo de bases de datos.
- OWL RL: Este sublenguaje permite ejecutar razonamientos sobre la ontología en tiempo polinomial utilizando técnicas de bases de datos extendidas por reglas aplicadas directamente a los datos representados como ternas RDF.

### 4. Componentes de OWL 2

Los componentes básicos definidos en OWL 2 para representar

el conocimiento son los siguientes:

- Axiomas: Son las verdades fundamentales del conocimiento que se modela en la ontología. Los axiomas definidos en una ontología son asumidos como conocimientos verdaderos.
- Entidades: Son los objetos que pertenecen al conocimiento que se está modelando. Las entidades pueden ser de varios tipos, como se verá más adelante.
- Expresiones: Son combinaciones de entidades que forman descripciones complejas de alguna parte del conocimiento a partir de las entidades básicas de éste.

#### 4.1 Tipos de Entidades

Las entidades del lenguaje OWL 2 pueden ser de tres tipos:

- Clases: Definen conceptos abstractos dentro del conocimiento que se modela. Entre las clases de una ontología se pueden definir relaciones de jerarquía, equivalencia, disyunción, entre otras.
- Propiedades: Las propiedades definen relaciones entre conceptos o adicionan información acerca de las clases definidas. Las propiedades, al igual que las clases, pueden ser organizadas jerárquicamente.
- Individuos: Son los objetos concretos del conocimiento que se modela y pertenecen a una o más clases. Nótese, sin embargo, que un mismo individuo no puede pertenecer a dos clases si éstas tienen una relación de disyunción entre ellas.

```
<owl:Class rdf:about="#AsparagusTopping">
  <rdfs:label xml:lang="pt"
  >CoberturaDeAspargos</rdfs:label>
  <rdfs:subClassOf>
    <owl:Class rdf:about="#VegetableTopping"/>
  </rdfs:subClassOf>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="#hasSpiciness"/>
      <owl:someValuesFrom rdf:resource="#Mild"/>
    </owl:Restriction>
  </rdfs:subClassOf>
</owl:Class>
```

Listado 1: Ejemplo de definiciones de clases, subclasses y propiedades en OWL 2 utilizando la sintaxis RDF/XML.

### 5. Aplicaciones

Desde la publicación de los primeros formatos para la representación del conocimiento por medio de ontologías en formatos digitales, son muchas las aplicaciones que se han encontrado a estas herramientas y cada vez se descubren más. Sin embargo, cabe destacar los siguientes aspectos de las Ciencias de la Información que hoy en día son centro de mucha atención por parte de los investigadores en este área.

#### 5.1 La Web Semántica

El concepto de Web Semántica es un esfuerzo liderado por el creador del World Wide Web, Sir Tim Berners-Lee, y la W3 el cual busca dotar a la red global de la capacidad de entender la información en ella contenida. Esto se pretende lograr adicionando metadatos a la información disponible en la red los cuales describan la naturaleza de la misma.

Es entonces evidente que OWL 2 ocupa un lugar prominente en

este esfuerzo [1], puesto que es un lenguaje especialmente diseñado para estructurar el conocimiento acerca de un tema de la realidad de manera que se pueda compartir fácilmente en la red.

La Web Semántica permitirá, por ejemplo, crear motores de búsquedas sensibles al contexto, de manera que los usuarios que busquen información relacionada con la palabra “carro” reciban también aquella relacionada con la palabra “automóvil”; o reconocer automáticamente cuál es el círculo de amigos de un usuario basándose en los metadatos disponibles del mismo.

A pesar de que el desarrollo de la Web Semántica ofrece un futuro prometedor a la computación como la conocemos hoy en día, existen enormes dificultades que hacen de este proyecto una de las tareas más colosales a la que se enfrenta la W3. En general, todos los retos que esta tarea impone están relacionados con el hecho de que la información es agregada a la red global de manera descentralizada, por lo tanto es imposible controlar de qué manera es ejecutado este proceso y no todos los usuarios de la red están alineados con la filosofía detrás de la Web Semántica. Para solventar estas dificultades, grandes esfuerzos se han realizado en torno a la creación automatizada de ontologías basada en datos estructurados ya presentes en lo que se conoce como la Red Profunda, la estructura de datos que soporta la Red pero que no está disponible a los usuarios [7].

## 5.2 Bases de Datos Inteligentes

Existen en la actualidad una gran cantidad de modelos de bases de datos, algunos más conocidos que otros, como lo son el Modelo Relacional y el Modelo Orientado a Objetos. Sin embargo, las bases de datos han llegado a ocupar un papel de tal importancia dentro de las aplicaciones que los Sistemas Manejadores de Bases de Datos (DBMS, por sus siglas en inglés) están siendo equipados con características que los hacen capaces de deducir nueva información acerca de la aplicación a partir de los datos ya almacenados. A un DBMS que cuenta con estas características se le conoce con el nombre de Sistema Inteligente Manejador de Bases de Datos (IDBMS), y este tipo de DBMS son cada vez más comunes, sobre todo entre las aplicaciones comerciales.

Las características que presente una IDBMS que la catalogan como tal pueden ser de distintos tipos. Algunos ejemplos de características inteligentes en una base de datos son la capacidad de editar datos almacenados de manera automática al reconocer ciertos estados dentro de la estructura de la información a punto de ser ingresada (esto puede ser implementado mediante la ejecución de reglas semánticas), o la capacidad de devolver al usuario información no almacenada en la base de datos pero que puede ser deducida a partir de los datos presentes en la misma. Sin embargo, una necesidad común que comparten todos estos sistemas es la necesidad de definir el conocimiento acerca de la información de manera estructurada, y es en este aspecto en el que los desarrolladores de estos sistemas utilizan las ontologías y los lenguajes para describirlas [8].

## 6. Programas y Librerías

Sin duda, la aceptación de un lenguaje para la definición de ontologías depende en gran manera de la disponibilidad de herramientas que faciliten la creación y edición de las mismas.

En esta sección presentamos algunas de las herramientas y librerías de programación más comunes que ayudan a trabajar con ontologías definidas en OWL 2.

### 6.1 Protégé

Esta plataforma de edición de ontologías hecha en Java es considerada el estándar de facto en su área [9]. La misma es desarrollada por la Universidad de Stanford con la colaboración de la Universidad de Manchester. Este software es libre y abierto, bajo la Mozilla Public License 1.1, por lo que ha amasado una enorme cantidad de colaboradores alrededor del mundo que conforman una gran comunidad de usuarios. Los mismos contribuyen módulos a la plataforma principal a través del Protégé API: una librería de programación en Java altamente documentada que permite acceder a las distintas áreas de la interfaz gráfica de la plataforma y a su funcionalidad interna.

La instalación básica de Protégé proporciona un conjunto estándar de componentes que incluyen visualizadores gráficos de relaciones (mediante el proyecto GraphViz [10]) y razonadores (delegando esta funcionalidad a librerías como Pellet y FaCT++, ver las siguientes secciones).

### 6.2 OWL API

Consiste en una interfaz de aplicación en Java que permite trabajar en la creación, edición y manejo de ontologías. Este proyecto, liderado por la Universidad de Manchester, es también abierto, bajo la licencia Lesser General Public License, y recibe grandes contribuciones de la comunidad de usuarios del mismo. Su desarrollo está estrechamente relacionado con el del proyecto Protégé (en especial mediante el Protégé-OWL API) sin embargo es importante destacar que no es necesario contar con la plataforma Protégé para utilizar el OWL API, por lo que este último puede ser incluido, por ejemplo, en el desarrollo de una aplicación web basada en Java que requiera manejar ontologías.

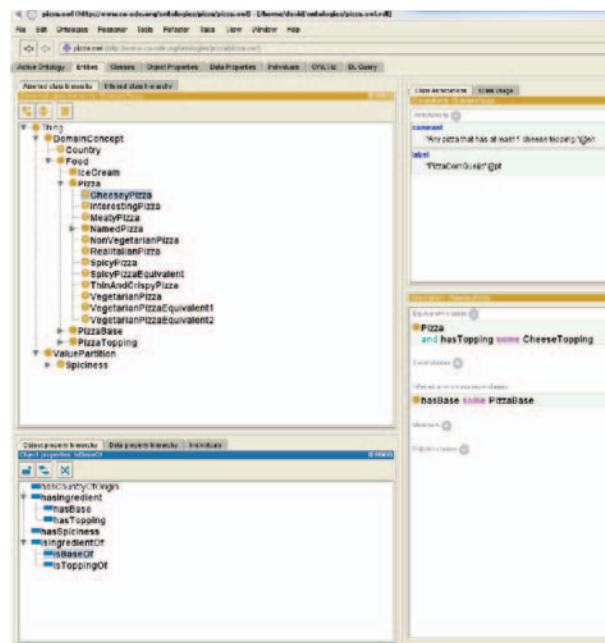


Figura 1: Vistazo de Protégé editando la ontología de la pizza.

Esta interfaz de programación cuenta con lectores y serializadores de documentos para las sintaxis RDF/XML, OWL/XML, OWL DL y Turtle, por lo tanto permite una enorme flexibilidad a la hora de escoger el formato para describir el conocimiento de la aplicación en la que se trabaje. Además de esto, el OWL API proporciona interfaces para razonadores, de manera que se pueda integrar con Pellet, FaCT++, entre otros.

Hasta el momento, el proyecto ha pasado por 3 versiones mayores, la última de ellas centrada en implementar todas las características de OWL 2.

### 6.3 Pellet

Pellet es un razonador implementado en Java que permite hacer inferencias y consultas basadas en ontologías descritas en OWL 2. Esta librería es desarrollada bajo un esquema de licenciamiento dual, donde el uso de la misma como parte de proyectos de Código Abierto está regido bajo los términos de la Affero General Public License. Este proyecto es patrocinado por la compañía Clark & Parsia LLC, la cual ofrece también soporte a usuarios comerciales.

Actualmente son muchos los usuarios de este razonador el cual es probablemente el más popular debido a su fácil integración con Protégé.

### 6.4 FaCT++

Este razonador que se distingue del resto de los paquetes de software y librerías antes mencionados en el hecho de que está implementado en C++. Este proyecto, bajo la Lesser General Public License, es una mejora sobre el razonador FaCT, un proyecto realizado por el Profesor Ian Horrocks de la Universidad de Oxford y sus colaboradores en 1998. FaCT++ utiliza los mismos algoritmos desarrollados en su predecesor pero con una arquitectura interna distinta.

### Referencias

- [1] Horrocks, Ian, "Ontologies and the semantic web", Communications of the ACM, vol. 51, pp 58--67, Diciembre 2008
- [2] Klyne, Graham and Carroll, Jeremy. (2004). Resource Description Framework (RDF): Concepts and Abstract Syntax. [Online]. Disponible en: <http://www.w3.org/TR/2004/REC-rdf-concepts-20040210/>
- [3] Brickley, Dan and Guha, R.V.. (2004). RDF Vocabulary Description Language 1.0: RDF Schema. [Online]. Disponible en: <http://www.w3.org/TR/2004/REC-rdf-schema-20040210/>
- [4] Schneider, Peter et. al. (2004). OWL Web Ontology Language Semantics and Abstract Syntax. [Online]. Disponible en: <http://www.w3.org/TR/2004/REC-owl-semantics-20040210/>
- [5] W3C OWL Working Group. (2009). OWL 2 Web Ontology Language Document Overview. [Online]. Disponible en: <http://www.w3.org/TR/2009/REC-owl2-overview-20091027/>
- [6] Motik, Boris et al. (2009). OWL 2 Web Ontology Language Profiles. [Online]. Disponible en: <http://www.w3.org/TR/2009/REC-owl2-profiles-20091027/>
- [7] An, Yoo Jung et al, "Semantic deep web: automatic attribute extraction from the deep web data sources", in Procs. of the 2007 ACM symposium on Applied comput, 2007, pp. 1667--1672
- [8] Muñoz, Ana et al, "A General Ontology for Intelligent Database", International Journal of Computers, vol. 1, pp 102--108, 2007
- [9] Knublauch,, "The Protégé OWL Plugin: An Open Development Environment for Semantic Web Applications", in The Semantic Web - ISWC 2004, 2004, pp. 229--243
- [10] Emden R. Gansner and Stephen C. North, "An Open Graph Visualization System and Its Applications to Software Engineering", Software - Practice and Experience, vol. 30, pp 1203--1233, 1999



## FACULTAD DE INGENIERÍA CIVIL

### Oferta Académica

La Universidad Tecnológica de Panamá, a la vanguardia con el incesante crecimiento de nuestro país; se ha destacado, durante los últimos años, por ser una de las universidades más sobresalientes y reconocidas internacionalmente por preparar óptimos y competentes profesionales en las áreas de ciencias y tecnología.

- Licenciatura en Ingeniería Agrícola.
- Licenciatura en Ingeniería Civil.
- Licenciatura en Ingeniería Geomática.
- Licenciatura en Ingeniería Ambiental.
- Licenciatura en Ingeniería Marítima y Portuaria.
- Licenciatura en Operaciones Marítimas y Portuarias.
- Licenciatura en Dibujo Automatizado / Título intermedio de Técnico en Ingeniería con especialización en Dibujo Automatizado.
- Licenciatura en Edificaciones / Título intermedio de Técnico en Ing. con espec. en Edificaciones.
- Licenciatura en Tecnología de Riego y Drenaje.
- Licenciatura en Saneamiento y Ambiente / Título intermedio de Técnico en Ing. con espec. en Saneamiento y Ambiente.
- Licenciatura en Topografía / Título intermedio de Técnico en Ing. con espec. en Topografía.
- Técnico en Ingeniería con especialización en Riego y Drenaje.