

Recibido: 05.02.2019 | Aceptado: 01.03.2019

Palabras clave: CERN, ontologías, semántica, web y w3c.

Web 3.0 = Web 2.0 + semántica



FRANCISCO EDGAR CASTILLO BARRERA
ecastillo@uaslp.mx
FACULTAD DE INGENIERÍA, UASLP

Es muy probable que los términos Web 1.0, Web 2.0, Web 3.0 y Web 4.0 no sean tan familiares en comparación con la palabra internet, creado antes que la Web, por la década de 1960 como resultado del proyecto Advanced Research Projects Agency Network (ARPANET) solicitado por el Departamento de Defensa (DOD) de Estados Unidos de América (EUA), para tener una red de computadoras que pudiera transferir información entre diferentes lugares con diferentes objetivos. Es una red mundial de comunicación en-



Figura 1.
Logo del CERN
y foto de Sir
Timothy John
Berners-Lee

tre computadoras. Años después, en el Centro Europeo para la Investigación Nuclear (CERN), Sir Timothy John Berners-Lee (TimBL) (figura 1), egresado de la Universidad de Oxford, Inglaterra, en 1990, en colaboración con Robert Cailliau, (figura 2), maestro en Ciencias de la Computación por la Universidad de Michigan, EUA, crearon la red de hipertexto mundial World Wide Web (www), que en su forma compacta se le conoce como web. También le llaman algunos actualmente la web de los datos enlazados o la web de datos.

Cabe mencionar que fue en el CERN donde se tuvo el primer servidor web en una computadora de la empresa NeXT, la NeXTcube, mejor conocida como The cube.

Para el funcionamiento de la web, ésta hace uso de la red internet para transferir datos por medio del protocolo de transferencia de hipertextos (HTTP) que maneja las conexiones entre páginas, cuenta con un localizador de recursos uniforme (URL, por sus siglas en inglés),



Figura 2.
Robert Cailliau (1995)
Imagen del CERN

cuyo objetivo es asignar una dirección única a cada uno de los recursos que forman parte de la red para su fácil localización (páginas, imágenes, etcétera) y, finalmente, las páginas web escritas en lenguaje de marcas de hipertexto (HTML, por sus siglas en inglés), el cual básicamente consiste en etiquetas llamadas *tags* que ayudan a describir los contenidos del texto y a los programas informáticos en la búsqueda de información en las páginas. Para entrar a la web se necesita tener conexión a internet y un navegador que soporte las tres tecnologías (HTML, HTTP y URL).

Para usar la red internet no es necesario un navegador, los dispositivos móviles se conectan y hacen uso de sus servicios; por ejemplo, el correo electrónico, información del clima o noticias. Actualmente se busca que todos los tipos de dispositivos puedan integrarse a la red sin necesidad de un navegador, a esto se le conoce como el internet de las cosas (IoT, por sus siglas en inglés).

Es importante destacar que para el crecimiento de la web existe una organización creada por TimBL en 1994, en el Instituto de Tecnología de Massachusetts (MIT) en conjunto con el CERN, de la cual él es su presidente: El Consorcio de la World Wide Web (W3C) (figura 3), tiene como función principal definir los estándares y las tecnologías para el futuro de ella.

¿Qué es la Web 3.0 o web semántica?

Como ya se mencionó, ha habido una evolución de la web, iniciando con la Web 1.0 que se caracterizaba por tener datos estáticos, es decir, no podían ser modificados por los usuarios, sólo por el encargado del sitio (conocido como webmaster), así que únicamente eran de lectura. Posteriormente, en 2004, Tom O' Reilly habla de la Web 2.0, a la que también algunos le llaman la web social, en donde los usuarios controlan sus contenidos, su información y con la aparición de las redes sociales se fomenta la interacción entre usuarios de diferentes partes del mundo. La Web 3.0 es una visión de TimBL por tener una red mundial para compartir información, basada en documentos estáticos y dotada de más significado o semántica, la cual le permitiría a las

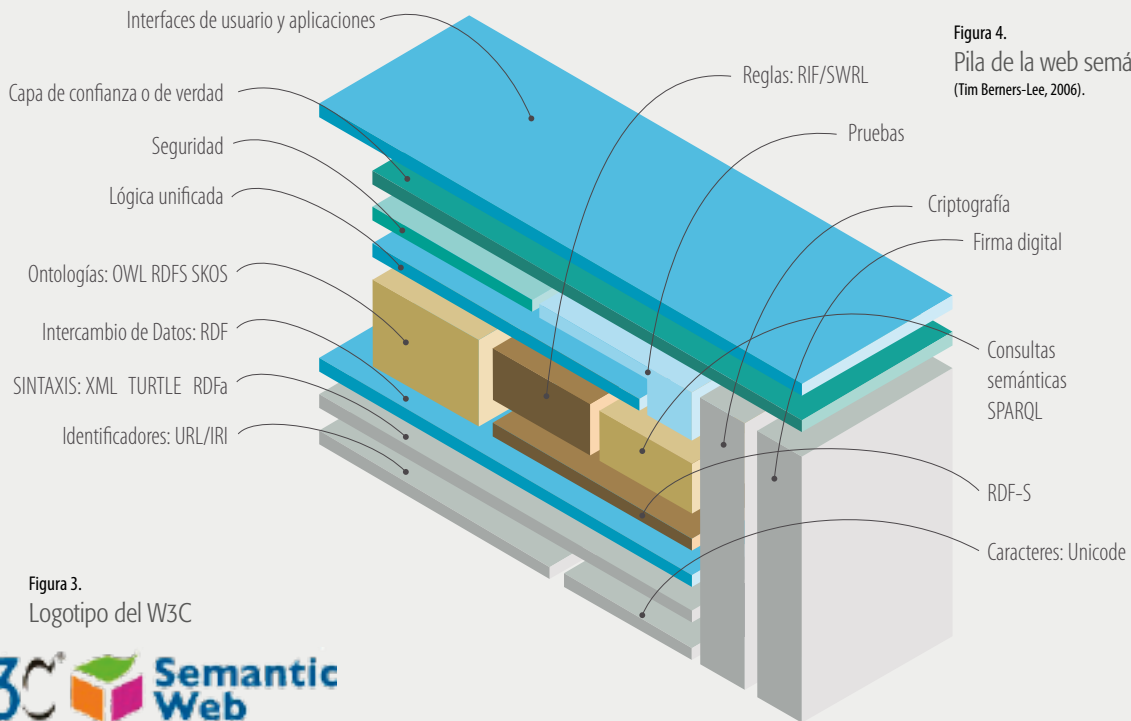


Figura 4.
Pila de la web semántica
(Tim Berners-Lee, 2006).

Figura 3.
Logotipo del W3C



máquinas y a los humanos entender la información contenida en ella, volviéndose algo más que una estructura sintáctica (lo que fue la Web 1.0).

El objetivo principal de la Web 3.0 es añadir contenido semántico y metadatos (datos que describen datos) a las páginas web para obtener información más exacta y precisa, en lugar de obtener demasiados links a sitios que no forman parte del dominio de lo que se busca. Tim Berners-Lee, James Hendler y Ora Lassila afirmaron en el año 2001 en el artículo "The Semantic Web" de la revista *Scientific American* que la web semántica no es una Web diferente, pero sí es una extensión de la actual, en la que la información está dada con un significado definido, permitiendo a las computadoras y a las personas trabajar conjuntamente.

Hay muchos retos por resolver y la inteligencia artificial juega un papel fundamental. Para lograrlos, TimBL propuso en el año 2000 una arquitectura por capas conocida como pila de la web semántica (Semantic Web Stack) o capas del pastel de la web semántica (Semantic Web Layer Cake), en la cual presentó las tecnologías que consideró necesarias: Las tecnologías de la web semántica son todas aquellas herramientas que proporcionan información semántica sobre el significado de las palabras y dan soporte a la web semántica (W3C, 2002) (figura 4).

En la figura 4 se muestran las capas de la pila de la web semántica en su tra-

ducción al español y con descripción de sus módulos.

Podemos clasificar las capas que conforman la web semántica en tres tipos de tecnologías: *a)* de hipertexto, *b)* estandarizadas y *c)* sin realizar.

Tecnologías web de hipertexto

Esta capa está conformada por las tecnologías conocidas como hipertexto¹, en informática es el texto que contiene enlaces que se asocian a otros, denominados hipervínculos, ligas o links, las cuales son la base del funcionamiento de la web. En este bloque tenemos:

- a)* Identificador de recursos internacionales (IRI, por sus siglas en inglés). Este tipo de hipertexto es una generalización del identificador de recursos uniforme (URI, por sus siglas en inglés), el cual define un único identificador para los recursos de la web semántica.
- b)* Unicode. Es un estándar de codificación de caracteres que asigna un número único para cada carácter. Su objetivo es facilitar la representación y manipulación de texto en diferentes lenguajes e incluso hasta en lenguas muertas. La web semántica necesita poder representar documentos en diferentes lenguajes usados por el ser humano.
- c)* Lenguaje de marcado extensible (XML, por sus siglas en inglés). Permite crear documentos compuestos de datos estructurados y definir etiquetas propias. La web semántica necesita juntar datos en un mismo documento, por tal motivo requiere

hacer referencia a varias fuentes en un sólo documento.

- d) Lenguaje terse RDF triple (Turtle). Es un formato que se usa para representar datos en forma de tripletas, cuyos elementos son un sujeto, un predicado y un objeto. En el ejemplo, el sujeto es "programación", el predicado es "es_profesor" y el objeto es "edgar". Se maneja por medio de los URI y para evitar escribirlo todo se usan prefijos, a los cuales se asigna un nombre corto, por ejemplo dc, uni, rdf, como se muestra en el ejemplo de la figura 5.
- e) Marco de descripción de recursos en atributos (RDFa). Es una extensión del lenguaje de marcado de hipertexto sensible (XHTML, por sus siglas en inglés), añadiendo más semántica que la proporcionada por el XHTML.

Tecnologías web estandarizadas

Las capas de enmedio están basadas en tecnologías que han sido estandarizadas por la W3C, las cuales permiten crear aplicaciones de la web semántica. Entre ellas tenemos:

“ El objetivo principal de la Web 3.0 es añadir contenido semántico y metadatos a las páginas web para obtener información más precisa

a) Marco de descripción de recursos (RDF, por sus siglas en inglés).

Es un marco de trabajo para especificar metadatos usando la notación de tripletas. Permite representar la información de los recursos por medio de un grafo, como se muestra en la figura 6. Hay que aclarar que se le llama recurso a cualquier cosa a la que pueda asignarse una URI, y un grafo en matemáticas y ciencias de la computación es un conjunto de nodos (vértices) y un conjunto de arcos (aristas). Se representan gráficamente con círculos (vértices) que contienen los datos y estos se conectan a otros por medio de líneas (aristas).

b) RDF Schema, RDF-S o RDFS proporciona un vocabulario básico para RDF.

c) Lenguaje de ontología web (OWL, por su acrónimo en inglés). Es un lenguaje de marcado para definir ontologías (su objetivo es identificar el significado correcto de las palabras polisémicas según el contexto de los documentos). Permite publicar y compartir los datos en la web. Tiene como objetivo facilitar un modelo de

Figura 5: Ejemplo de notación en Turtle.

Notación Turtle

```
@prefix dc:      <http://purl.org/dc/elements/1.1/>.
@prefix uni:    <http://uaslp.mx/>

uni:programación  dc:es_profesor  uni:edgar
```

Tripeta RDF

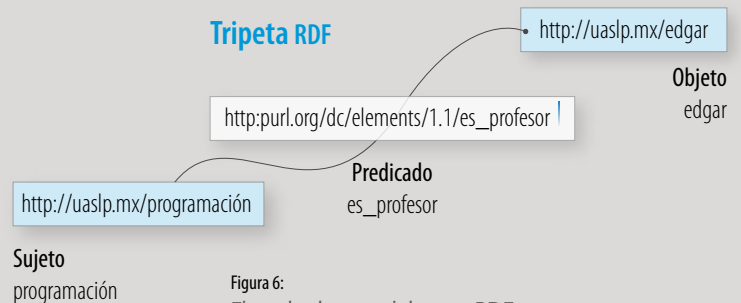


Figura 6: Ejemplo de una tripleta en RDF.



**FRANCISCO
EDGAR CASTILLO
BARRERA**

Es doctor en Tecnologías de la Información por la Universidad de Guadalajara. Actualmente es profesor investigador en el Área de Ciencias de la Computación de la UASLP y trabaja en el proyecto "OntBot: Sistema Interactivo en Lenguaje Natural basado en Ontologías".



marcado construido sobre RDF y codificado en XML. Está basado en la lógica descriptiva, la cual le proporciona a la web la capacidad de razonamiento (modelos formales que le permiten hacer inferencia sobre los datos para deducir información), la cual es sin duda el inicio de una web con inteligencia artificial.

d) Protocolo y lenguaje de consulta RDF (SPARQL, por su acrónimo en inglés). Es un lenguaje de consultas para RDF; permite que las aplicaciones de la web semántica puedan obtener información de las ontologías.

Tecnologías web sin realizar


Las últimas capas o capas superiores todavía no han sido estandarizadas y, en algunos casos, sólo se tienen ideas de cómo deben ser implementadas para lograr la web semántica. Los módulos que forman parte de esta última capa son:

- a) Lenguaje de reglas de web semántica (SWRL).** Es un lenguaje que permite definir reglas de producción y proporciona una solución al querer describir propiedades o acciones que no puedan expresarse usando la lógica descriptiva en OWL.
- b) Formato de intercambio de reglas (RIF).** Es una recomendación del W3C y es parte de la infraestructura para la web semántica.
- c) Lógica unificada.** Con ella se busca hacer inferencia sobre los datos.
- d) Pruebas (Proof).** Se busca tener un lenguaje unificado que permita obtener inferencias lógicas por medio del uso de reglas. Para ello es necesario tener un único identificador asociado a los recursos.

e) Capa de confianza (Trust). Es la capa encargada de validar los contenidos, la cual proporciona dos formas para obtener o derivar sentencias que sean correctas por medio de verificar que las premisas vengan de una fuente confiable y se aplique lógica formal durante la inferencia de nueva información.

f) Interfaces de usuario y aplicaciones. Es la capa final que permite a los humanos utilizar aplicaciones de la web semántica.

Finalmente, en forma paralela a todos los módulos y capas, es necesario tener una firma digital que garantice la autenticidad e integridad de los datos obtenidos y el uso de la criptografía para ocultar aquellos que son confidenciales y viajan por la red.

Una firma digital es una forma de validar la originalidad, autoría e integridad de los datos de un documento que es enviado por la red, siendo tan válida como los documentos que se firman físicamente; algunas aplicaciones manejan el concepto de certificado digital. Cada software comercial tiene su forma de creación, la más usada es por técnicas y algoritmos matemáticos que toman los datos que proporciona la persona que firma (sólo ella los conoce) y los transforma. 

Referencias bibliográficas:

- Antoniou, G., Grout, P., Van Harmelen, F. y Hoekstra, R. (2004). *A Semantic Web Primer*. Massachusetts: Massachusetts Institute of Technology.
- Berners-Lee, T., Hendler, J. y Lassila, O. (2001). The Semantic Web. *Scientific American*, 284(5), pp. 28-37.
- Labra Gayo, J. E. (2011). *Web semántica: comprendiendo el cambio hacia la Web 3.0*. Coruña: Netbiblo.