



Ciencias
de la
Computación

Ingeniantes

Un enfoque para la gestión de contenido semántico en aplicaciones del Internet de las cosas (IoT)



Colaboración

Alan Omar Urías Osuna, Ricardo Quintero Meza, Luis Carlos Santillán, Emir Manjarrez, Instituto Tecnológico de Culiacán

RESUMEN: Actualmente el Internet de las Cosas (IoT) es un paradigma que integra varias tecnologías, con el fin de interconectar objetos cotidianos a través de Internet; por otro lado, la Web Semántica se conoce por una infraestructura eficaz para mejorar la visibilidad de la información de la web. Entre el IoT y la Web Semántica existe distanciamiento, debido a que el IoT es un paradigma el cual requiere de muchos esfuerzos y se le ha dado mayor importancia a la implementación que a la estructuración del conocimiento extraído. Por este motivo, en este artículo, se propone la definición de un enfoque basado en una arquitectura de referencia que permite gestionar el contenido semántico con información extraída de aplicaciones IoT, de esta manera se obtiene un enriquecimiento de la información. El enfoque es validado a partir de un caso de estudio orientado al dominio de monitoreo de transportes.

PALABRAS CLAVE: Internet de las Cosas (IoT), Ontología, Web Semántica.

ABSTRACT: Currently the Internet of Things (IoT) is a paradigm that integrates various technologies, with the purpose of interconnect with everyday objects through the Internet; on the other hand, the semantic web, is known for an effective infrastructure to improve the visibility of information on the web. There is distancing between the IoT and the semantic web, since the IoT is a paradigm which requires many efforts and has been given greater importance to the implementation than the structuring of the extracted knowledge. For this reason, in this article, we propose the definition of an approach based on a reference architecture that allows to manage the semantic content with information extracted from IoT applications, and in this way obtained an enrichment of the information. The approach is validated from a case study oriented on transport monitoring domain.

KEYWORDS: Internet of Things (IoT), Ontology, Web Semantic.

I. INTRODUCCIÓN

El Internet de las Cosas [1] es un concepto que se refiere a la interconexión digital de objetos cotidianos con Internet. En la actualidad el Internet de las Cosas (IoT) ha surgido con mucha fuerza, arrasando con la ola digital que está aplicándose en diferentes áreas como lo son los medios de publicidad para tener con exactitud la correcta segmentación del mercado; el medio ambiente, con sensores para la optimización de recursos; industrias aumentando la calidad de servicios y reduciendo costos; manufactura, estableciendo redes de control y procesos de fabricación; energía, optimizando el consumo de energía; medicina, utilizando notificación de emergencia, monitores de presión arterial y frecuencia cardiaca; transporte, implementando un control de tráfico inteligente, control de vehículos y sobre todo seguridad para los usuarios.

Por otro lado, la Web Semántica [2] es una extensión de la web actual, en la cual los datos se describen y relacionan de tal manera que establecen por si solos un contexto o significado, lo cual no solo les permite a los usuarios encontrar datos, sino que también otras aplicaciones se pueden integrar y entender la información fácilmente. Es por ello la importancia de implementar ontologías que hagan uso de vocabularios estándares para que un mayor número de aplicaciones puedan entender los contextos y dominios gracias a estos esquemas definidos en las ontologías.

Debido al incremento y crecimiento en desarrollo de aplicaciones IoT en diferentes áreas, éstas producen información generada por los objetos incluidos en esas aplicaciones debido a sensores. Los sensores generan una gran cantidad de datos crudos, de los cuales apoyados de técnicas de Web Semántica se pueden dotar de sentido para enriquecer su conocimiento.

Una de las técnicas importantes para la representación de la información en la Web Semántica son las ontologías. Las ontologías definen los términos y las relaciones básicas para la comprensión de un área del conocimiento, utilizando reglas para poder combinar términos, también facilitan un vocabulario controlado que comparte conceptos en común con un dominio, para representar la información disponible a consultar, además las ontologías permiten definir relaciones entre dichos conceptos. Tanto conceptos como roles pueden ser usados para realizar consultas más complejas y recuperar de forma precisa la información en la que el usuario está interesado en consultar.

La Web Semántica y el paradigma del Internet de las Cosas (IoT) no ha tenido un acercamiento profundo con la extracción de conocimiento de la Web Semántica, debido a la complejidad del IoT se han enfocado más en la implementación (sensores) y se ha dejado la extracción del conocimiento en una menor prioridad, por lo que se propone una arquitectura para la extracción de esta información y dotarla de conocimiento para su óptimo uso.

Actualmente, la integración basada en ontologías es un área de investigación muy activa en la Web Semántica. Es por esto que se diseñó de una arquitectura de referencia para gestionar la información obtenida de aplicaciones IoT de manera semántica. El objetivo de esto es obtener información de aplicaciones del IoT, representar esa información con técnicas de Web Semántica para que los usuarios tengan datos más enriquecidos del dominio.

La información recolectada por los sensores del sistema de monitoreo de transporte escolar podría ser requerida por diferentes interesados:

- Padres de familia,
- Instituciones educativas,
- Autoridades policíacas y viales, etc.

Por lo tanto, en este artículo se propone diseñar una arquitectura de referencia que permita la gestión de contenido semántico orientado a aplicaciones de IoT, que permitan la integración de datos enriquecidos y esquemas estandarizados para que puedan ser utilizados o consumidos por otras aplicaciones orientadas al IoT; de esta manera se permite a usuarios tales como: particulares, tutores, autoridades viales, instituciones educativas, dueños de transportes y más, consultar dicha información basada en ontologías diseñadas específicamente para el dominio que se especifique. Para este artículo se enfocó en el dominio de monitoreo de transportes, dando la ventaja de obtener la información que requieren los usuarios antes mencionados. Para validar el enfoque propuesto en este artículo, se presenta un ejemplo orientado al dominio de un sistema de monitoreo de transportes escolares [3].

La arquitectura propuesta permite a usuarios interesados consultar la información semántica que proveen las aplicaciones IoT, mediante de la ontología creada para este propósito. Para la construcción de la arquitectura de referencia se utilizaron algunos componentes open source como: Apache Jena, IDE Eclipse, protégé, Pallet, etc.

Este artículo está estructurado de la siguiente manera: En la sección I se presenta la introducción; En la sección II se presentan los trabajos relacionados; En la sección III se presenta el enfoque para la gestión de contenido semántico, que contiene la metodología a seguir comenzando por el diseño de la ontología, posteriormente definir una arquitectura propuesta que se descompone en diagrama de contexto, componentes, pruebas y finalmente las tecnologías utilizadas para poder implementar dicha arquitectura; En la sección IV se presentan las conclusiones de este artículo.

II. TRABAJOS RELACIONADOS.

En esta sección se mencionan algunos trabajos relacionados con el monitoreo de transportes, ontologías e Internet de las Cosas (IoT) que atañen a este artículo. A continuación se describen algunos de ellos.

A.Towards Semantic Monitoring Data Collection and Representation in Federated Infrastructures.

En este trabajo desarrollaron un modelo basado en ontologías para el intercambio interoperable de datos a través de múltiples dominios administrativos con el fin de lograr que la heterogeneidad de las herramientas e interfaces utilizadas. Este trabajo presenta un enfoque para la unificación de la recolección de datos de vigilancia y la representación basada en las tecnologías de la Web Semántica y su principal contribución es el diseño e implementación de un modelo de información basado en ontologías para el monitoreo de las infraestructuras federadas, la solución prevista se realizó en el contexto del proyecto europeo Fed4Fire [4].

B. Interoperability of Security-Enabled Internet of Things.

En este trabajo se propone una arquitectura en capas para un “framework” orientado al Internet de las Cosas enriquecido semánticamente, conectando entre sí las otras capas y facilitar la provisión de acceso seguro a los servicios habilitados del Internet de las Cosas. El principal aporte es asegurar el razonamiento entre las ontologías y reglas semánticas, también la interoperabilidad de los aspectos de seguridad se aborda mediante la ontología y una plataforma de máquina a máquina. Este trabajo proporciona detalles de implementación del razonamiento de seguridad y los aspectos de interoperabilidad y discute los desafíos cruciales entre estas áreas [5].

C. On a Java based implementation of ontology evolution processes based on Natural Language Processing.

En este trabajo se realizó una implementación basada en Java donde el aporte principal fue la evolución del proceso de llevar el lenguaje natural hacia una ontología, con este software se extrajo información estructurada de lenguaje natural y se utilizó para enriquecer la ontología. Se cumplieron los objetivos de extracción de conocimiento legible de texto y transformar la información extraída en una forma que es adecuada para su inserción en una ontología [6].

D. Incorporating technology in service-oriented i* business models: a case study.

En este trabajo se presenta un enfoque para la incorporación de tecnología en los modelos de negocio, lo que permite evaluar sistemáticamente la contribución de una tecnología específica para cumplir con los objetivos de negocio. Una de las ventajas de este enfoque es que tanto software como hardware son representados de una manera abstracta. Por lo tanto, los nuevos componentes de software o nuevos componentes de hardware se pueden representar por lo que sus propiedades explícitas como atributos de calidad. Otra ventaja es la representación explícita de las relaciones entre los objetivos de la organización y las actividades de bajo nivel necesarios para realizar los procesos [7].

III. ENFOQUE PARA GESTIÓN DE CONTENIDO SEMÁNTICO

Para poder aplicar un enfoque para gestión de contenido semántico es importante realizar el diseño de una ontología, como técnica de la Web Semántica forma parte medular en este enfoque debido que es donde se dota de conocimiento a los datos; por otro lado y no menos importante se encuentra el diseño de una arquitectura propuesta en este caso, debido a que se debe definir la misma para poder ser implementada y tomada como referencia para cualquier otro enfoque que intersecte la Web Semántica y se puedan implementar diferentes ontologías en diferentes dominios enfocados al IoT.

A. DISEÑO DE LA ONTOLOGÍA

Las ontologías [8], son colecciones de enunciados escritos en un lenguaje, como el RDF, que define las relaciones entre conceptos y especifica reglas lógicas para razonar con ellos; Una ontología es una especificación explícita y formal de una conceptualización compartida. Una conceptualización es una vista simplificada y abstracta del mundo que deseamos representar para algún propósito en específico, definiendo un vocabulario controlado.

Las ontologías están formadas de diversos componentes como lo son Conceptos, Relaciones, Funciones, Instancias y Axiomas.

Las ontologías definen un conjunto de clases, relaciones, funciones y constantes para un determinado dominio, además introducen axiomas para restringir la interpretación de estos elementos. Las relaciones que se establecen entre los elementos del dominio son aquellas que clasifican las entidades del dominio en base a relaciones. Para el caso de estudio se siguió la metodología Methontology [9], para la creación de la ontología, y para la creación se utilizó la herramienta Protégé [10].

En las ontologías todo parte de una “cosa”, definiendo clases a las cuales se les define propiedades y atributos correspondientes; Como parte del caso de estudio de monitoreo de transportes, se definió la siguiente ontología, con las clases de monitoreo, vehículo, localización, conductor e información, que se ilustra parte de ella en la Figura 1.

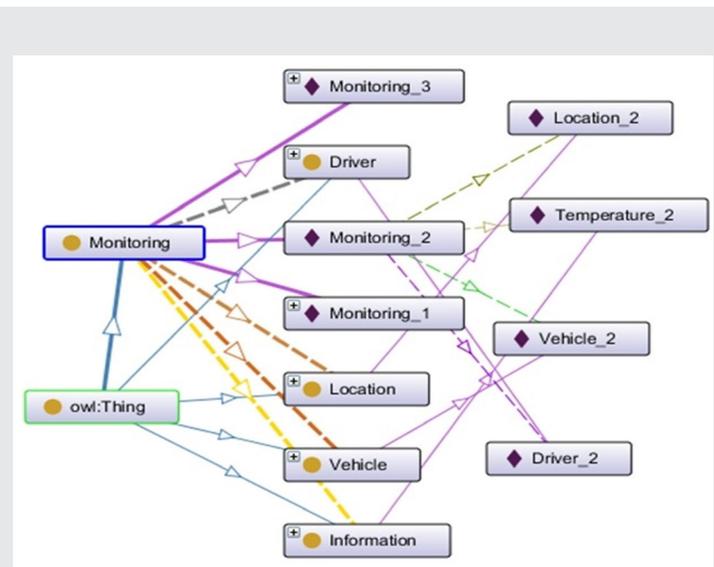


Figura 1. Parte de ontología orientada al monitoreo de transportes creada en herramienta Protégé.

B. ARQUITECTURA PROPUESTA

En esta sección, se detalla la arquitectura para un enfoque de gestión de contenido semántico basado en IoT con dominio en monitoreo de transportes.

Para definir la arquitectura, es necesario establecer los límites del sistema con entidades externas, para ello

se utiliza el diagrama de contexto el cual se detalla a continuación:

1. Diagrama de contexto

En la Figura 2. Se muestra el diagrama de contexto, donde se propone que el sistema “Monitoreo Ontológico de Transportes” sea la pieza central; las interfaces 1 y 2 son aplicaciones externas que requieren consumir los servicios del sistema, en este caso los interesados en obtener la información son aplicaciones de los padres de familia e instituciones educativas que desean saber en todo momento el nivel de seguridad con los que los estudiantes son transportados; las autoridades policíacas vigilan que el chofer no exceda los límites de velocidad, representados en la Interface 3.

En el diagrama de contexto se muestra una relación al mismo nivel de colaboración entre el sistema Monitoreo Ontológico Escolar y la aplicación IoT Bus representada en la interface 4, la cual colecta información a través de sensores basados en IoT, esta información es integrada en la base de datos de SQL del sistema de monitoreo ontológico mostrado en interface 6; El sistema representado como “Monitoreo Ontológico de Transportes” tiene una dependencia directa para su óptima funcionalidad con una base de datos de triples para que los usuarios puedan explotar la información que ellos requieren consultar, representado en la interface 5.

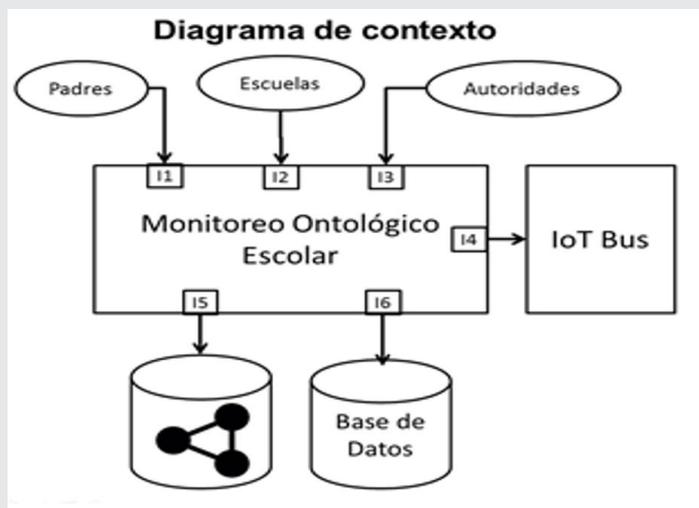


Figura 2. Diagrama de contexto.

2. Componentes

Conocer los componentes externos hacia los cuales el sistema de Monitoreo Ontológico de Transportes tiene dependencia, es importante al momento de seleccionar un estilo arquitectónico y definir la estructura del sistema en términos de componentes.

Por la naturaleza de la aplicación se decidió implementar un estilo arquitectónico basado en capas. Cada capa presta un servicio específico y permite disminuir las dependencias entre los diferentes componentes.

A continuación, se describen las capas y su responsabilidad:

a. SDB: Base de Datos en SQL.

Este componente es el encargado de facilitar la comunicación del sistema con la base de datos, tanto para integrar los datos colectados por IoT Bus, como para facilitar la consulta de los mismos desde la ontología web.

b. TDB: Base de Datos de Triples.

Este componente es el encargado de facilitar la comunicación del sistema con la base de datos de triples o RDF Store, tanto para integrar los datos colectados por IoT Bus, como para facilitar la consulta de los mismos desde la ontología web.

c. Ontology Engine.

Este componente es el encargado de implementar la ontología, la cual es definida en archivos que contienen los modelos creados en formato RDF/XML y OWL (Ontology Web Language) en los cuales se definen las reglas de inferencia y estructura de la ontología orientada al monitoreo de transportes generando el contenido semántico de la información. También se encarga de ejecutar las consultas que se solicitan a la ontología en lenguaje SPARQL.

Este motor ontológico se implementa en Java extendiendo los frameworks de apache jena y pellet.

d. WebHub.

Este componente permite recibir peticiones vía web services en formato SPARQL, valida la petición del usuario y solicita al motor ontológico realizar la consulta y retornar los resultados obtenidos.

e. Integration:

Este componente permite obtener los datos del sistema IoT Bus que deposita los datos en el componente 1 SDB (base de datos en SQL).

En la tabla 1. Se muestra la relación entre modelo de contexto y estructura de componentes.

Tabla 1: Relación entre modelo de contexto y estructura de componentes.

Interface	Componente
I1 (Padres)	4 - WebHub
I2 (Escuelas)	4 - WebHub
I3 (Autoridades)	4 - WebHub
I4 (Bus IoT)	5 - Integration
I5 RDF Store	2 - TDB
I6 SQL DB	1 - SDB

3. Pruebas

Para probar la factibilidad de la arquitectura propuesta se emplea un diagrama de instancias, conciliando con el diagrama de contexto y componentes donde se validará que todos los requisitos sean cubiertos por la arquitectura propuesta, este diagrama de instancias se ilustra en la Figura 3.

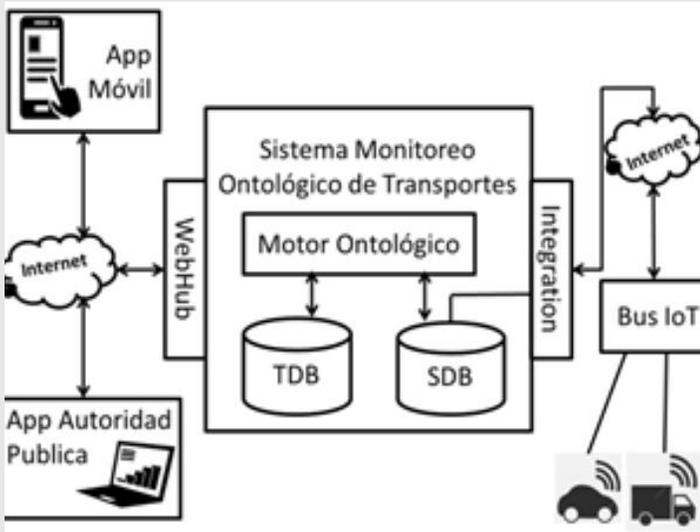


Figura 3. Diagrama de Instancias.

En la Figura 4, se muestra la arquitectura propuesta, en la cual se tiene como parte central un motor ontológico, que tiene como entrada algunos archivos que contienen los modelos creados en formato RDF/XML y OWL (Ontology Web Language), en los cuales se crearon las reglas de inferencia y estructura de la ontología orientada al monitoreo de transportes; se propone una aplicación para que los usuarios finales puedan realizar las consultas de la información de monitoreo de transportes en la cual es proporcionada por el framework “Bus IoT” mencionado en el diagrama de contexto, que este a su vez deja la información en la capa de persistencia “Store API - SDB” que es una base de datos en SQL; en la capa de persistencia “Store API - TDB” tenemos almacenados los “Triples” que nos sirven para estructurar la información que es extraída de la base de datos en SQL, que aportan a la información para convertirla en contenido semántico posteriormente.

La capa de “WebHub” es un servicio web que son consumidos por las consultas desde cualquier dispositivo disponible para la aplicación. Finalmente centrándose en la capa del motor ontológico, este dota de conocimiento a la información extraída convirtiéndola en contenido semántico por medio de una ontología bien definida con esquemas estándares, que es enfocada al monitoreo de transportes, por lo que al momento de consultar se cumpla el objetivo que es satisfacer las necesidades de los usuarios que requieren respuestas precisas con respecto a la información buscada.

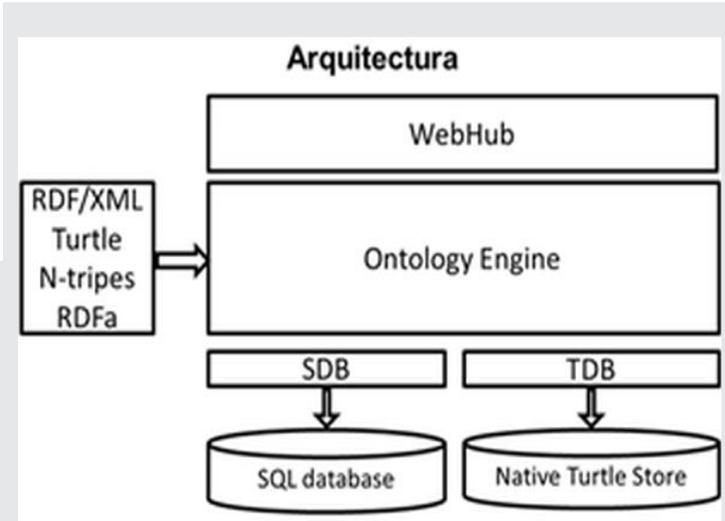


Figura 4. Arquitectura propuesta para el enfoque de gestión de conocimiento semántico en aplicaciones IoT en el dominio de monitoreo de transportes.

4. Tecnología

Se utilizaron varios componentes tecnológicos que expondrán a continuación:

- Apache Jena que proporciona un componente que es útil para este proyecto, “Apache Jena TDB”, este componente sirve como almacenamiento y consultas de RDF, el cual utiliza otro componente en la capa de “WebHub”.
- El componente en la capa de “WebHub” (detallado en la arquitectura) debe proporcionar un servidor SPARQL donde se pueden utilizar TDB para el almacenamiento de persistencia y proporcionar los protocolos de SPARQL para consultar, actualizar y actualización de REST (Representational State Transfer) a través de HTTP.
- En la parte de ontologías se utilizó una herramienta llamada “Protégé” la cual es una plataforma libre, de código abierto que proporciona un conjunto de herramientas para la construcción de modelos de dominio y aplicaciones basadas en ontologías, dotando a estas mismas de conocimiento.
- También se utilizó la tecnología de IDE Eclipse Standard/SDK Versión: Kepler Service Release 2 Build id: 20140224-0627, para el desarrollo del motor ontológico y la integración de los componentes.
- Se utilizó a su vez “Pellet Reasoner” que es un razonador de OWL (Ontology Web Language) Pellet se puede utilizar con Jena o bibliotecas OWL-API.
- Pellet proporciona funcionalidad para comprobar la coherencia de las ontologías, calcular la jerarquía de clasificación, especifica inferencias y responder a las consultas en SPARQL.
- Por otra parte se utilizaron vocabularios RDF estándar [1] para vehículos, localización y personas, con el fin de que cualquier aplicación que maneje este vocabulario RDF pueda razonar con dicha información y de esta forma se puedan integrar diversas aplicaciones y consultar la información disponible dotada de conocimiento mediante la Web Semántica.

Se crearon archivos en formato OWL para definir la semántica de la ontología, del cual se muestra un fragmento de código en la Figura 5.

```
ontology-33;Monitoring"/>
  <hasDate rdf:datatype="&xsd;dateTimeStamp">2016-04-21
  11:22:33.444</hasDate>
  <isDriverBy rdf:resource="&untitled-ontology-33;Driver_
  3"/>
  <isLocation rdf:resource="&untitled-
  ontology-33;Location_3"/>
  <isInformation rdf:resource="&untitled-
  ontology-33;Temperature_3"/>
  <isVehicle rdf:resource="&untitled-ontology-33;Vehicle_
  3"/>
  </owl:NamedIndividual>

  <!--
  http://www.semanticweb.org/alan/ontologies/2016/3/untitled-
  ontology-33#Temperature_1 -->

  <owl:NamedIndividual rdf:about="&untitled-
  ontology-33;Temperature_1">
    <rdf:type rdf:resource="&untitled-
  ontology-33;Information"/>
    <hasValue rdf:datatype="&xsd;float">35.0</hasValue>
    <owl:sameAs rdf:resource="&untitled-
  ontology-33;Temperature_2"/>
    <owl:sameAs rdf:resource="&untitled-
  ontology-33;Temperature_3"/>
  </owl:NamedIndividual>
```

Figura 5 Fragmento de código realizado en OWL.

Se crearon archivos en formato RDF del cual se muestra un fragmento de código en la Figura 6.

```
<rdf:RDF
  xmlns="http://www.semanticweb.org/alan/ontologies/2016/3/untitled-ontology-33#"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  >
  <rdf:Description
  rdf:about="http://www.semanticweb.org/alan/ontologies/2016/3/untitled-ontology-33#Temperature_1">
    <rdf:type
  rdf:resource="http://www.semanticweb.org/alan/ontologies/2016/3/untitled-ontology-33#Information"/>
    <hasValue
  rdf:datatype="http://www.w3.org/2001/XMLSchema#float">35.0</hasValue>
    <rdf:type
  rdf:resource="http://www.w3.org/2002/07/owl#NamedIndividual"/>
  </rdf:Description>
  <rdf:Description
  rdf:about="http://www.semanticweb.org/alan/ontologies/2016/3/untitled-ontology-33#LocationTmp">
    <hasLatitude
  rdf:datatype="http://www.w3.org/2001/XMLSchema#float">24.7885</hasLatitude>
    <rdf:type
  rdf:resource="http://www.semanticweb.org/alan/ontologies/2016/3/untitled-ontology-33#Location"/>
    <rdf:type
  rdf:resource="http://www.w3.org/2002/07/owl#NamedIndividual"/>
```

Figura 6 Fragmento de código realizado en OWL.

IV. CONCLUSIONES

Este trabajo a diferencia de los relacionados anteriormente, tiene la ventaja de contar con una arquitectura propuesta para desarrollar diferentes aplicaciones y sus componentes necesarios enfocados al monitoreo de transportes en general orientados en IoT.

También se cuenta con una ontología diseñada específicamente para el monitoreo de transportes escolares,

donde podemos expandir a placer dependiendo de la diversidad de información que se quiere monitorear, teniendo diferentes sensores para generar esta información y reutilizar esta ontología para adaptarse a cualquier otra aplicación de monitoreo de transportes ya que cuenta con esquemas estándares.

Gracias a estos esquemas estándares que la hacen compatible, puede ser usada por otro tipo de aplicaciones de IoT orientadas al monitoreo de transportes ya que empatarán en el dominio, por ejemplo, monitoreo de transporte de alimentos, monitoreo de transporte de vehículos (Nodrizas), donde se tendría que adaptar la ontología para aceptar algún elemento no tomado en cuenta.

En base a las validaciones realizadas, podemos concluir que es posible la construcción de una arquitectura de referencia que facilite la integración de diferentes aplicaciones por medio de la web; es muy viable desarrollar ontologías para dominios específicos y utilizar vocabularios RDF estándar para la interconexión o interoperabilidad de aplicaciones, enfocándose en el intercambio de información y que puedan razonar el contenido semántico de la información disponible entre diferentes aplicaciones, haciéndose así extensiones unas de otras.

En este caso probamos con el monitoreo de transportes escolares, pero hay muchos dominios que pueden ser explorados y desarrollar ontologías estándares para construir un modelo del conocimiento basado en Web Semántica, con esto concluimos que en un futuro no muy lejano podremos obtener el mayor aprovechamiento de la web 3.0 dotando de conocimiento semántico a toda (o por lo menos la mayor parte de) la información que tenemos en la web.

BIBLIOGRAFÍA

[1] Bahga, A., & Madisetti, V. (2014). *Internet of Things: A Hands-on Approach*. VPT.

[2] Tim Berners-Lee, James Hendler, Ora Lassila, *The Semantic Web*, *Scientific American*, May 2001.

[3] J. Zambada, R. Quintero, R. Isijara, R. Galeana, and L. Santillan, "An IoT based scholar bus monitoring system," *IEEE International Smart Cities Conference*, 2015.

[4] Al-Hazmi, Y., & Magedanz, T. (2015, August). *Towards Semantic Monitoring Data Collection and Representation in Federated Infrastructures*. In *Future Internet of Things and Cloud (FiCloud)*, 2015 3rd International Conference on (pp. 17-24). IEEE.

[5] Alam, S., Chowdhury, M. M., & Noll, J. (2011). *Interoperability of security-enabled Internet of things*. *Wireless Personal Communications*, 61(3), 567-586.

[6] Gabbanini, F. (2010). *On a Java based implementation of ontology evolution processes based on Natural Language Processing*.

[7] Martinez, A., Vazquez, B., Estrada, H., Santillan, L., & Zavala, C. *Incorporating technology in service-oriented i* business models: a case study*. *Information Systems and e-Business Management*, 1-27.

[8] Maedche, A., & Staab, S. (2001). *Ontology learning for the semantic web*. *IEEE Intelligent systems*, 16(2), 72-79.

[9] Jones, D., Bench-Capon, T., & Visser, P. (1998). *Methodologies for ontology development*.

[10] Stanford Research, 'Protégé', June 2015. [Online]. Available: <http://protege.stanford.edu/> URL. [Accessed: 19-Ene-2017].

[11] Schema, 'Home - schema.org', April 2015. [Online]. Available: <http://schema.org/> URL. [Accessed: 19-Ene-2017].

