

Universidad Autónoma Metropolitana Unidad Azcapotzalco
División de Ciencias Básicas e Ingeniería
Ingeniería en Computación

Proyecto de Investigación

"Gestión de Horarios de Cursos Mediante Ontologías"

Trimestre 2016 Otoño

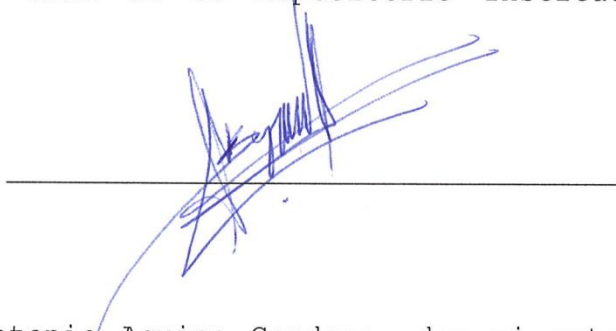
Irving Antonio Aquino Cordova
207330482

Asesora:
Maricela Claudia Bravo Contreras
Profesor Asociado
Departamento de Sistemas

9 de Enero del 2017

Declaratoria

Yo, Maricela Claudia Bravo Contreras, declaro que aprobé el contenido del presente Reporte de Proyecto de Integración y doy mi autorización para su publicación en la Biblioteca Digital, así como en el Repositorio Institucional de UAM Azcapotzalco.



Yo, Irving Antonio Aquino Cordova, doy mi autorización a la Coordinación de Servicios de Información de la Universidad Autónoma Metropolitana, Unidad Azcapotzalco, para publicar el presente documento en la Biblioteca Digital, así como en el Repositorio Institucional de UAM Azcapotzalco.



Resumen

Este proyecto consistió en la automatización de la población de una ontología, en la cual se gestionan los horarios de la Universidad Autónoma Metropolitana de la Unidad Azcapotzalco.

En principio se diseñaron y crearon tres diferentes ontologías nombradas: Espacio Físico, UEA y Persona, con atributos únicos de cada una, y se importaron dentro de otra ontología llamada Horario, donde se maneja todas las relaciones entre cada una de ellas.

Todo esto realizo a partir de un programa en lenguaje Java, el cual exporta una hoja de cálculo a la ontología Horario.

Tabla de contenido

1.	Introducción	7
2.	Antecedentes	8
3.	Justificación	10
4.	Objetivos	11
	4.1. Objetivo general	11
	4.2. Objetivos específicos	11
5.	Marco teórico	12
	5.1. Ontología	12
6.	Desarrollo del proyecto	13
	6.1. Ontología Persona	13
	6.2. Ontología UEA	14
	6.3. Ontología Espacio Físico	14
	6.4. Ontología Horario	15
	6.5. Programa Java de auto-poblado	16
	6.5.1. Botón Agregar Archivo	17
	6.5.2. Botones para Agregar Ontologías	17
7.	Resultados	18
8.	Conclusiones	20
9.	Referencias bibliográficas	21

Índice de Figuras

Figura 1. Diagrama de Ontología Persona	13
Figura 2. Diagrama de Ontología UEA	14
Figura 3. Diagrama de Ontología Espacio Físico	14
Figura 4. Diagrama de Ontología Horario	15
Figura 5. Aplicación con un Archivo Cargado	18
Figura 6. Ejemplo de Poblado de Ontología Persona	18
Figura 7 Ontología Persona Vacía	19
Figura 8 Ontología Persona Poblada	19

Índice de Tablas

Tabla 1. Data Properties de la Ontología Horario	16
Tabla 2. Object Properties de la Ontología Horario	16

1. Introducción

Las ontologías son modelos que permiten la representación de datos con atributos cambiantes. En este documento se propone el desarrollo de un proyecto tecnológico para la simplificación de la representación y manejo de horarios de cursos.

Actualmente, en la División de Ciencias Básicas e Ingeniería de la UAM Azcapotzalco, la información referente a los horarios de clases se gestiona a través de plantillas de hojas de cálculo, con la ayuda de la nueva aplicación que se planea realizar esta información se extraería desde estas tablas a un sistema basado en ontologías que facilite la consulta de dicha información, con esto se espera que todo se encuentre mejor organizado para una búsqueda más clara y mejor clasificada.

2. Antecedentes

Proyectos Terminales

1. Sistema de búsqueda y recuperación de ontologías [1]

Este proyecto consiste en una aplicación web que realiza búsquedas y recuperación de ontologías a partir de palabras clave, el cual será de ayuda como referencia para el sistema que se quiere programar en cuanto a la parte de consulta que se realizara para los horarios.

2. Asistente para la evaluación de Ontologías [2]

Esta aplicación web tiene la función de evaluar ontologías, con la finalidad de analizar que cumpla con los principios de diseño; se encontró esto de utilidad para la parte que realizara la verificación de la ontología, una vez importada la información.

3. Sistema Web para el Poblado Automático de Ontologías a partir de Textos [3]

Este sistema extrae información necesaria para poblar ontologías automáticamente desde archivos de texto, en el caso de este proyecto en vez de ser archivos de texto, serán plantillas de hojas de cálculo.

Página Web

1. Plantilla de control de horarios en Excel [4]

Este sistema maneja los horarios por medio de plantillas en hojas de cálculo, a comparación del proyecto que tomara esta plantilla para una forma más simple de representación con ayuda de un sistema que maneje ontologías.

Articulo

1. Lenguaje de ontologías web aplicadas al sector turístico [5]

Este artículo explica como una búsqueda a base de ontologías funciona mucho mejor para generar resultados que son más exactos en cuanto a los criterios que el usuario necesita, es por esto que se quiere utilizar este tipo de búsqueda para la consulta de horarios.

2. Uso de ontologías y web semántica para apoyar la gestión del conocimiento [6]

El artículo explica cómo se utilizan de manera provechosa los esquemas de búsquedas de sistemas Web Semánticos para catalogar de forma más organizada la gran cantidad de información que se manejan, y de misma forma se pretende aprovechar esto para un mejor funcionamiento del sistema.

3. Justificación

La ventaja de las ontologías, a comparación de una base de datos, es que dentro de las ontologías se permite la definición de consultas para extraer información utilizando las múltiples dimensiones y cualquier número de atributos; esto es más práctico a la hora de consultar y modificar horarios.

Además al ser un modelo basado en lenguajes formales y lógica descriptiva, permite la descripción de clases, jerarquías de clases, así como las relaciones semánticas entre conceptos y axiomas; esto favorece a la reutilización, mantenimiento, la actualización y expansión (modularización).

Si se integran todas estas ventajas en un sistema de gestión de horarios resulta bastante útil, ya que al manejar grandes cantidades de datos, será más fácil extraer y utilizar la información desde el punto de vista del usuario.

4. Objetivo

4.1. Objetivo general

Diseñar e implementar un sistema que genere de manera automática una ontología de horarios a partir de plantillas en hojas de cálculo con la información de los cursos que se planea impartir cada trimestre en el Departamento de Sistemas para la carrera de Ingeniería en Computación.

4.2. Objetivos específicos

1. Diseñar y construir el modelo de una ontología de horarios considerando la importación y reutilización de tres ontologías que representan a Personas, Programas de Estudio y Espacio Físico, previamente diseñadas.
2. Diseñar e implementar un sistema que permita analizar sintácticamente y extraer los datos de los cursos programados desde la plantilla de una hoja de cálculo para que éstos se almacenen en la ontología.
3. Diseñar e implementar los módulos que permitan realizar la verificación de la consistencia lógica entre el modelo diseñado y las instancias almacenadas en la ontología; esto con el propósito de mantener la ontología consistente.

5. Marco Teórico

5.1. Ontología

Las ontologías definen de forma estándar y consensuada un vocabulario de conceptos así como las relaciones entre ellos dentro de área concreta del conocimiento, formando redes jerárquicas semánticas. Actúan a modo de tesauros, pero en lenguaje informático. Así, recogen reglas lógicas y restricciones para hacer "comprender" a las máquinas los conceptos que manejan dentro de un determinado campo.

Las ontologías permiten trabajar con conceptos, en lugar de palabras clave, en los sistemas de recuperación de información. Desde el punto de vista de las fuentes de información, éstas describen el contenido de los repositorios de datos independientemente de la representación sintáctica de los mismos, posibilitando su integración semántica.

En el contexto del Web Semántico, las ontologías suponen una capa por encima de RDF y los esquemas RDF.

"Una ontología es una especificación explícita de una conceptualización", según Thomas Gruber. Una conceptualización es una abstracción, una vista simplificada del mundo que queremos representar. Una ontología no es una base de datos ni un programa (porque tienen sus propios formatos internos), no es una conceptualización (porque no es una especificación, es sólo una vista) ni una tabla de contenidos (aunque una taxonomía sí es una ontología).

Las ontologías son acuerdos, en un contexto social, para cubrir una serie de objetivos. Se crean, entre otras cosas, para:

- Permitir el intercambio de datos entre programas.
- Simplificar la unificación (o traducción) de distintas representaciones.
- Facilitar la comunicación entre personas

6. Desarrollo del proyecto

Se realizaron 3 ontologías: Espacio físico, UEA y Persona. Las cuales poseen diferentes atributos, así como clases y subclases, esto para un mejor manejo al momento de agruparlas en otra ontología más completa, donde se manejara el proyecto, dentro de esta hay atributos que relacionan a las otras tres.

6.1. Ontología Persona

Usando el programa Protegé se realizó el diseño de las ontologías. En la ontología persona (Figura 1), se jerarquizo a los académicos para llegar a la clase profesor, la cual usaría más adelante.



Figura 1. Diagrama de Ontología Persona

Dentro de esta Ontología solo utilice los data Properties que utilizaría para el manejo de horarios, los cuales son:

- tieneNombreProfesor
- tieneNumeroEconomico

Era lo único relevante para la clase profesor, ya que es lo único que se refleja en la hoja de cálculo de horarios y son los únicos atributos que se necesitan para este proyecto.

6.2. Ontología UEA

Del mismo modo que para la ontología Persona, se hizo el diseño de UEA (figura 2).



Figura 2 Diagrama de Ontología UEA

La diferencia es que solo se necesitó una clase, ya que el manejo de UEA's no necesito de otras clases para llegar a ella, dentro de los data Properties se usaron los que se requieren para su gestión; los cuales son:

- tieneNombreUEA
- tieneCreditos
- tieneClave

Estas son los atributos que necesitare para referenciar a las UEA's que se usaran.

6.3. Ontología Espacio Físico

Esta es la última ontología que se crea, para el manejo de horarios, ya se tiene los profesores, las UEA's, por lo que solo falta el lugar donde se impartirán, el diseño solo se compone del salón o sala que se usara (Figura 3).

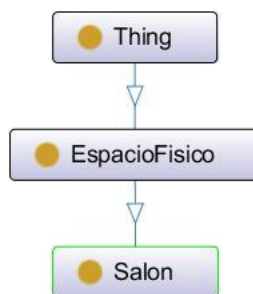


Figura 3 Diagrama de Ontología Espacio Físico

Dentro de los horarios solo se requiere saber el id con el que se conoce, cada salón por lo que solo se creó la siguiente data Property:

- tieneIdSalon

También dentro del identificador del salón se reconoce en que edificio se encuentra, por lo que no fue necesario agregarlo.

6.4. Ontología Horario

Por último se diseñó la ontología en la cual se importarían las tres anteriores la cual se llama Horario (Figura 4), en esta se crean object Properties para relacionar las clases entre sí; además de que se utilizan los data Properties de las demás ontologías.

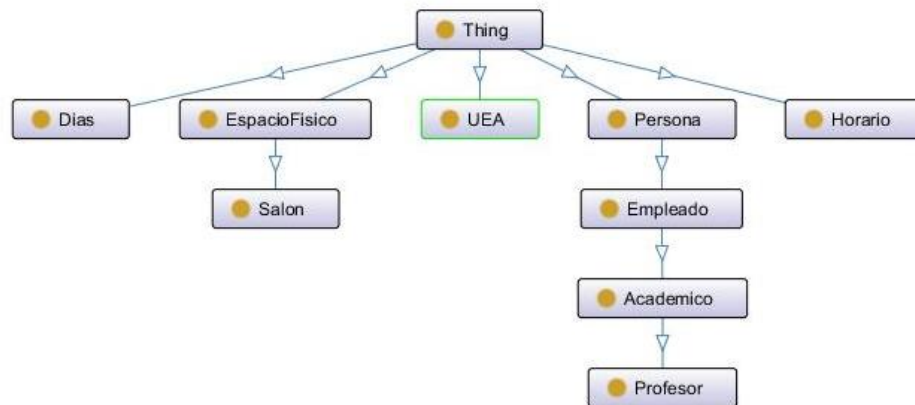


Figura 4 Diagrama de Ontología Horario

Aquí se usan todas las data properties de las demás ontologías y otras creadas para el uso de esta ontología. Todas las data properties que se usan son las siguientes:

Ontología	Data Property	Tipo
Persona	tieneNombreProfesor	String
Persona	tieneNumeroEconomico	Integer
UEA	tieneNombreUEA	String
UEA	tieneCreditos	Integer
UEA	tieneClave	Integer
EspacioFisico	tieneIDSalon	String
Horario	TieneCupo	Int
Horario	tieneHoraInicio	String
Horario	tieneHoraFin	String
Horario	tieneGrupo	String

Tabla 1 Data Properties de la Ontología Horario

De igual manera se crearon Object Properties para la ontología, las cuales ayudaran al manejo de los horarios.

Ontología	Object Property	Dominio	Rango
Horario	seImparteDias	UEA	Días
Horario	seImparteEn	UEA	Salón
Horario	seImpartePor	UEA	Profesor

Tabla 2 Object Properties de la Ontología Horario

6.5. Programa Java de auto-poblado

El lenguaje que se había propuesto desde un principio fue Java, ya que posee librerías para el manejo de ontologías, y que anteriormente ya había trabajado con ellas; así también, la librería para el manejo de hojas de cálculo. El programa que utilice para la creación de la interfaz fue Netbeans.

Lo que se propuso fue una interfaz gráfica sencilla para el usuario, en el cual solo hay cuatro botones:

- Agregar Archivo
- Agregar Ontologías Personas
- Agregar Ontologías UEA
- Agregar Ontologías Espacio Físico

6.5.1. Botón Agregar Archivo

En esta parte del código cree una jTable dentro de una interfaz gráfica Java, en la cual se manejó por vectores, para que cada que leyera una celda de la hoja de cálculo, creara una nueva celda, y cada que saltara a otra fila se pudiera crear una nueva en la interfaz, para la importación de los datos utilice la librería "org.apache.poi", me fue de mucha utilidad, ya que la pude utilizar para el manejo de archivos con terminación ".xlsx".

Básicamente lo que realiza es leer celda por celda la hoja de cálculo, extraer el contenido y pasarlo al jTable, es un proceso largo, pero es la única forma en que se lea todo el contenido.

6.5.2. Botones para Agregar Ontologías

Una vez creada la jTable con todos los datos, cree botones diferentes, para que importara a su respectiva ontología la información que necesitaba, en cada uno de ellos el proceso era el mismo, el cual consistía en leer los datos de persona, UEA's y Espacio Físico, almacenarlos en una lista de arreglos y eliminar los duplicados, para esto use las funciones Map y Hash Map, en los datos de Personas y UEA's; esto fue porque tenían varias Data Properties, de diferente tipo, cadena y entero.

En el caso de los profesores utilice los números económicos, ya que son variables que no pueden tener dos profesores el mismo, por esta misma razón use la clave de la UEA en su respectiva ontología, y después de crear una lista nueva sin repeticiones, creaba los individuos en la ontología que le correspondía y sus Data Properties.

En Espacio Físico fue más sencillo, ya que ahí solo se lee el identificador de cada salón, ya que ninguno tiene el mismo; agregarlo a una lista y para evitar las repeticiones use la función Hash Set. Al eliminar repeticiones la asigne a su ontología, creando individuos con su identificador correspondiente.

La librería que utilice para el manejo de la ontología fue "org.semanticweb.owlapi", la cual ayuda al proceso de poblar la ontología.

Resultados

Al iniciar la interfaz se debe presionar el botón agregar archivo, que cargara la información de la hoja de cálculo, la interfaz final, muestra la tabla que se importa (Figura 5)

U.E.A.	NOM.	CR.	Grupo	CUPO	LUNE	LUNE	MART	MART	MIER	MIER	JUEV	JUEV	VIER	VIER	SALON	ECO	PROF
1100...	Taller...	6	CDIV...	30												33448	MER...
1100...	Taller...	6	CDIV...	30												18618	ORT...
1100...	Taller...	6	CDIV...	30													
1100...	Taller...	6	CDIV81	30													
1100...	Taller...	6	CDIV82	30													
1100...	Taller...	6	CDIV83	30													
1100...	Taller...	6	CDIV84	30													
1151...	Elem...	9	CSI01	30	8:30	10:00			8:30	10:00			8:30	10:00			
1151...	Siste...	9	CSI01	40	11:30	13:00			11:30	13:00			11:30	13:00		22003	ESTR...
1151...	Siste...	9	CSI81	40	14:30	16:00			14:30	16:00			14:30	16:00			
1151...	Audit...	9	CSI01	30			12:15	14:30			12:15	14:30				15177	SAN...
1151...	Progr...	9	CSAC...	50												17114	SILVA...
1151...	Progr...	7	CSAL...	40	7:00	8:30			7:00	8:30			7:00	8:30		14937	GUTI...
1151...	Progr...	7	CTG01	40	8:30	10:00			8:30	10:00			8:30	10:00		27499	RAN...
1151...	Progr...	7	CTG02	40	0	0			0	0			0	0		16153	LAUR...
1151...	Progr...	7	CTG03	40	11:30	13:00			11:30	13:00			11:30	13:00		16153	LAUR...
1151...	Progr...	7	CTG04	40			0	0			0	0					
1151...	Progr...	7	CTG05	40			0	0			0	0				22664	KHAT...

Figura 5 Aplicación con un archivo cargado

Una vez cargada, se procede a usar los otros tres botones, no importa el orden es que se presionen, cada una cargara una ontología diferente y en consola se muestra el progreso de la población (Figura 6).

```

Output - ProyectoHorarios (run) x
ZAVALA OSORIO YADIRA
<http://www.semanticweb.org/ontologies/2016/4/Persona#Profesor>
<http://www.semanticweb.org/ontologies/2016/4/Persona#prof18>
LARQUE SAAVEDRA MARIO ULISES
<http://www.semanticweb.org/ontologies/2016/4/Persona#Profesor>
<http://www.semanticweb.org/ontologies/2016/4/Persona#prof19>
CORTES LEON HECTOR
<http://www.semanticweb.org/ontologies/2016/4/Persona#Profesor>
<http://www.semanticweb.org/ontologies/2016/4/Persona#prof20>
RAMIREZ RODRIGUEZ JAVIER
<http://www.semanticweb.org/ontologies/2016/4/Persona#Profesor>
<http://www.semanticweb.org/ontologies/2016/4/Persona#prof21>
GONZALEZ BELTRAN BEATRIZ ADRIANA
<http://www.semanticweb.org/ontologies/2016/4/Persona#Profesor>
  
```

Figura 6 Ejemplo de Poblado de Ontología Persona

Y para confirmar que se llenará adecuadamente se abrió la ontología en el programa Protegé.

En el siguiente ejemplo se muestra la ontología vacía (Figura 7) y después de ejecutarse el programa, se notan los individuos que se crearon con sus respectivas propiedades (Figura 8).

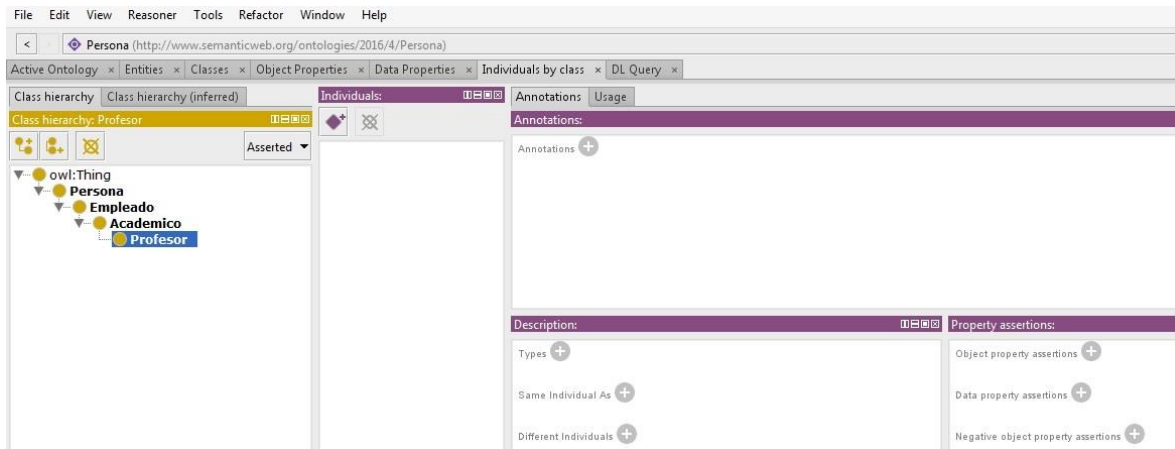


Figura 8 Ontología Persona Vacía

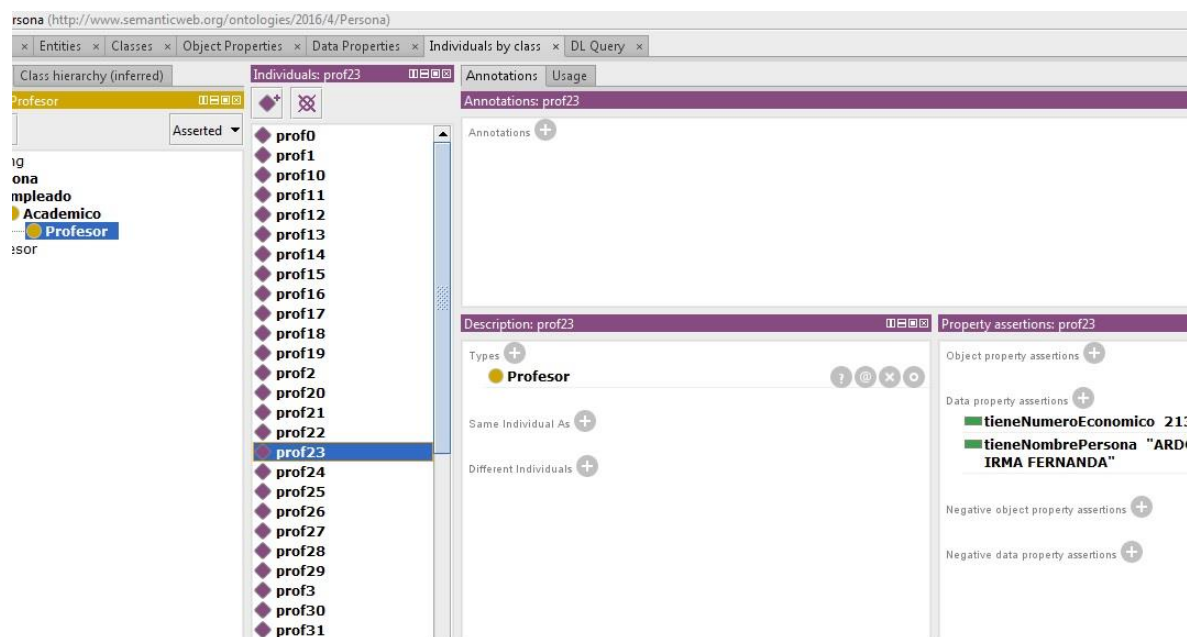


Figura 7 Ontología Persona Poblada

Conclusiones

Se cumplió el objetivo principal, el cual era el poblado automático de la ontología, así evitar llenarla de manera manual, lo cual se hace exhaustivo a toda la información que se maneja para organizar los horarios.

Una de los primeros problemas que encontré es que no todas las celdas tienen datos, ya que no todas las clases se imparten diario, así que son celdas que se quedan vacías y que no se pueden leer, ya que de hacerlo se leería todas las celdas del archivo, y eso tendría el programa corriendo infinitamente hasta que encuentre la última celda. Por lo que la solución fue llenar esas celdas con un carácter que no se usara, en mi caso use "_" (guion bajo), así cada que leía ese carácter, el programa sabía que debía pasarla aunque no tuviera contenido importante para cumplir su objetivo.

El poblado y manejo de la ontología Horario se queda como trabajo a futuro, ya que mi asesora me aconsejo que de ahí puede hacerse un nuevo proyecto, ya que todavía le falta mucho para que trabaje de una manera más óptima y con el fin de obtener mejores resultados, tanto a la hora de realizarlos, como al momento de hacer consultas.

Por el último motivo, no se han agregado manual de instalación, ni de uso; ya que este proyecto aun continuara.

Referencias Bibliográfica

- [1] A. C. Alamillo Medina, "Sistema de búsqueda y recuperación de ontologías", Proyecto Terminal, División Ciencias Básicas e Ingeniería, Universidad Autónoma Metropolitana, Unidad Azcapotzalco, 2013.
- [2] J. F. Santa Cruz Mora, "Asistente para la evaluación de Ontologías", Proyecto Terminal, División Ciencias Básicas e Ingeniería, Universidad Autónoma Metropolitana, Unidad Azcapotzalco, 2009.
- [3] C. M. Pilapanta Herrera, "Sistema Web para el Poblado Automático de Ontologías a partir de Textos", Proyecto Terminal, División Ciencias Básicas e Ingeniería, Universidad Autónoma Metropolitana, Unidad Azcapotzalco, 2015.
- [4] Excelworld en español, "Plantilla de control de horarios en Excel | Plantilla para descargar", 2016. [Online]. Available: <http://es.excelworld.net/plantillas/control-de-horarios-en-excel>. [Accessed: 01- Mar- 2016].
- [5] A. Soza and L. Garrido, "Lenguaje de ontologías web aplicadas al sector turístico", Revista Educación en Ingeniería, vol. 4, no. 8, pp. 122-129, 2009.
- [6] A. Cantor, "Uso de ontologías y web semántica para apoyar la gestión del conocimiento", Ciencia e Ingeniería Neogranadina, vol. 17, no. 2, pp. 111-129, 2007.
- [7] 2017. [Online]. Available: <https://www.infor.uva.es/~sblanco/Tesis/Ontolog%C3%ADas.pdf>. [Accessed: 04- Dec- 2016].