



Ontología Turística Geográfica

Memoria del Proyecto

Por David Socías Segura

TFC en el área de XML y Web Semántica
Enginyeria Tècnica en Informàtica de Sistemes

Consultor: Joan Casas Roma

Índice

Índice de contenido

1.Introducción.....	4
1.1.Descripción del proyecto.....	4
1.2.Objetivos del proyecto.....	4
1.3. Alcance.....	5
2.Introducción a la web semántica.....	6
2.1.Evolución de la Web: del origen de Web a la Web Semántica.....	6
2.2.Ontologías.....	7
2.3.XML y XML Schema.....	8
2.4.RDF y RDF-Schema.....	9
Las sentencias en RDF.....	9
RDF-Schema como herramienta para la creación de ontologías.....	9
Clases, propiedades e instancias.....	10
Dominios y rangos.....	10
Inferencia de nueva información a partir de información semántica.....	10
2.5.OWL.....	11
3.Recursos de interés turístico.....	13
3.1.Buscando información sobre puntos de interés turísticos	13
3.2.El problema de la información fragmentada.....	13
3.3.Una ontología turística	13
3.4.Objetivos de la ontología.....	14
3.5.Recursos digitales de interés turístico	14
Formato de un recurso digital.....	14
3.6.Puntos de interés turístico.....	15
Una visión geográfica.....	15
4.Ontología turística geográfica.....	16
4.1.Recursos semánticos en Internet	16
Una ontología en varios documentos.....	17
4.2.Puntos de Interés: Space.owl.....	17
Estructura general de las clases.....	18
Espacios construidos por el ser humano: Place.....	18
Espacios naturales: Geographical_feature.....	20
Estructura política y zonal: Geopolitical_entity.....	20
Definiciones exclusivas y no exclusivas.....	21
4.3.Recursos digitales: RecursosDigitales.owl.....	21
4.4.Ontología turística geográfica: RecursosTurisiticos.owl.....	22
5.Buscador de recursos turísticos.....	24
5.1.Requerimientos de la aplicación.....	24
5.2.Análisis de funcionalidades.....	24
5.3.Diseño de la aplicación.....	25
5.4.Interfaz de la aplicación.....	25
API para Web Semántica.....	27
Clases descriptivas.....	27
Controlador de gestión.....	28

Flujo de una consulta.....	29
5.5.Arquitectura del prototipo.....	30
Tecnología empleada.....	30
Estructura de capas.....	30
5.6.Instalación.....	31
Requerimientos de la aplicación	31
Instalación a partir del .WAR.....	31
Construir la aplicación a partir de las fuentes de la aplicación.....	31
6.Conclusiones.....	33
6.1.Conclusiones sobre la ontología creada.....	33
6.2.Consideraciones sobre el prototipo de buscador.....	34
7.Glosario de términos.....	36
8.Bibliografía.....	38
8.1.Libros.....	38
8.2.Páginas Web.....	38
9.Anexos.....	39
9.1.Anexo I. Enunciado del proyecto.....	39

1. Introducción

1.1. Descripción del proyecto

Este documento es la memoria del Trabajo Final de Carrera con el título "Ontología Turística Geográfica", donde se describe completamente el proyecto creado.

El proyecto trata sobre la creación, usando tecnologías de web semántica, de una ontología de carácter turístico que permita clasificar y relacionar documentos digitales en internet con puntos de interés turístico, y la creación de un buscador web que permita hacer consultas sobre un grupo de datos que utilicen la ontología creada. En este documento se explica qué es lo que entendemos por web semántica, que entendemos por ontología, recurso digital o punto de interés.

Este proyecto consta de cuatro partes, diferenciadas en este documento por capítulos, que abordan las diferentes fases en la creación de la aplicación final. Si tomamos como objetivo final la creación de la aplicación web pueden entenderse estos capítulos como las fases necesarias para llevar el proyecto a su conclusión.

El capítulo de Introducción a la Web Semántica muestra los resultados obtenidos de la fase de investigación previa, necesaria en cualquier proyecto que aborde el desarrollo de trabajos que contienen tecnología inicialmente desconocida. En este capítulo se hace un resumen de qué es lo que se entiende por Web Semántica, qué es una ontología, y cuáles son y cómo funcionan las tecnologías que emplearemos durante el desarrollo del proyecto.

El capítulo sobre Recursos de Interés turístico estudia el dominio sobre el que se va a trabajar. En él se definen los conceptos de punto de interés, lugar geográfico y recurso digital, y se observan las relaciones que existen entre ellos y que deberemos tener en cuenta a la hora de desarrollar el proyecto.

En el capítulo sobre Ontología Turística Geográfica se analiza y se diseña la ontología, cumpliendo el primero de los objetivos del proyecto. El capítulo profundiza en los motivos por los cuales se diseña la ontología de determinada manera, y se explican los problemas y las soluciones encontradas.

Por último, en el capítulo sobre el Buscador de Recursos Turísticos, se aborda un proyecto completo de desarrollo de software que, aunque limitado en tamaño, cumple con todas las fases de la ingeniería del software (análisis, diseño, desarrollo, pruebas e implantación).

Estos cuatro capítulos forman el tronco de este proyecto. El resto de capítulos aportan información necesaria en un documento de estas características, mostrando esta presentación del proyecto, las conclusiones finales, la biografía utilizada, y un glosario de términos.

1.2. Objetivos del proyecto

Este proyecto tiene dos objetivos principales bien diferenciados, que son la creación de una ontología de ámbito turístico y la creación de un prototipo web que permita

consultar la ontología creada. A partir de estos objetivos principales, vamos a definir los objetivos particulares de cada uno de ellos.

Los objetivos de la ontología son:

- Ha de modelar recursos digitales.
- Ha de modelar puntos de interés turístico.
- Ha de permitir georreferenciar los puntos de interés.
- Ha de enlazar recursos digitales con los puntos de interés que muestra
- Hay que crear unas instancias utilizando la ontología para usarlas como banco de pruebas del proyecto.

Los objetivos del prototipo de buscador son:

- Ha de acceder a la ontología y las instancias creadas para demostrar su uso.
- Ha de permitir buscar puntos de interés por nombre, tipo, lugar y recursos digitales asociados.
- Ha de mostrar fichas de los puntos de interés donde se vean los recursos digitales asociados y un mapa con su localización.

1.3. Alcance

El desarrollo completo de este proyecto incluye la creación de una ontología turística geográfica, la creación de un conjunto de pruebas en base a instancias de la ontología creada, la creación de un prototipo de buscador web que haga uso de esta ontología y sus instancias, la creación de un vídeo de presentación del proyecto y la creación de este documento de memoria. En este punto se explica el alcance de cada uno de estos productos, indicando el punto de vista con el que han sido creados.

La ontología ha sido creada con el objetivo de que sea utilizable en multitud de tipos de proyecto con mayor o menor relación con este. La ontología incluye clases y propiedades que no son utilizadas en este proyecto, pero que cumplen con este alcance global para permitir que sea utilizada en muchos proyectos con pocas o ninguna modificación.

El conjunto de instancias de prueba se ha creado pensando en demostrar algunas capacidades de la ontología más que en que sea una base extensa. De este modo sólo hay una veintena de instancias de puntos de interés y algo más del doble de recursos turísticos. El ámbito geográfico es el de Mallorca, y aún así, sólo hay unos pocos municipios representados. A pesar de su poca extensión, el grupo de pruebas demuestra la potencia de la ontología incluyendo ejemplos que muestran los puntos fuertes de nuestra ontología.

El prototipo de buscador es simple. Se ha construido de este modo para no complicar el uso ya que su objetivo es permitir navegar por las instancias de pruebas, comprobando las ventajas de la ontología. A la hora de usarlo, el buscador sigue un proceso basado en búsqueda, lista de resultados y ficha de punto de interés.

El vídeo de presentación, presenta la ontología y el buscador haciendo énfasis en su

uso más que en su desarrollo. Podría considerarse que se ha construido a modo de manual de uso, mostrando cómo se usa tanto la ontología como el buscador.

Este documento de memoria está pensado como resumen de todos los procesos que dan como resultado el proyecto. En este documento se explica el qué, el cómo y el porqué de los productos del proyecto.

2. Introducción a la web semántica

Comprender todos los detalles de la tecnología de la web semántica ocuparía varios libros. No es el cometido de este proyecto ahondar en estos detalles. Aún así, es necesario un conocimiento básico de la historia, los objetivos y la estructuración de la información en esta tecnología para comprender el desarrollo de este proyecto.

En este capítulo se explica a grandes rasgos qué se entiende por web semántica, y cómo se estructura el conocimiento en ella. Son unas pocas pinceladas con el objetivo de explicar varios conceptos clave. Para profundizar en el conocimiento de esta tecnología y en la sintaxis de RDF, RDFS y OWL, se recomienda una lectura de los textos indicados en la bibliografía.

2.1. Evolución de la Web: del origen de Web a la Web Semántica

Se considera que la World Wide Web como tal fue creada por Sir Tim Berners-Lee en 1989. Cansado de problemas de compatibilidad de formatos y de permisos de acceso a la hora de publicar documentos con sus colegas en el CERN, Sir Tim Berners-Lee escribió una propuesta a sus superiores que acabaría convirtiéndose en lo que es la Red actual.

En sus inicios, la web fue diseñada como una forma simple de acceder a documentos situados en diversos lugares sin tener que preocuparse por permisos de acceso o formatos.

El HTML permitía esta independencia de formatos, y la web se construyó a partir de documentos estáticos que contenían información mezclada con formatos de maquetación.

Escribir y preparar catálogos estáticos es costoso, mucho más si el origen de esos catálogos es dinámico y cambia cada poco tiempo. Para evitar ese trabajo de constante actualización empezaron a surgir aplicaciones que generaban los HTML de forma dinámica. De este modo surgieron los CGI y posteriormente los lenguajes de Script, que permitían generar dinámicamente las páginas web a partir de las bases de datos, y también el camino inverso, introduciendo datos en bases de datos a partir de formularios rellenados por el usuario en una página web.

El auge de las webs dinámicas expandió la utilidad de la red. Es posible encontrar infinidad de catálogos dinámicos en internet. Las empresas y organizaciones podían mostrar sus productos y actuaciones a sus clientes, informándoles sobre lo que la empresa tenía disponible para ellos. Pero esta comunicación existía en un solo sentido, los usuarios eran espectadores mudos de toda esta información.

Como evolución a esto, empezaron a surgir webs que permitían al usuario

expresarse, cerrando el círculo de comunicación entre los proveedores y los consumidores y diluyendo las diferencias entre estos. De este modo los usuarios, antes mudos, pueden votar, comentar, proponer y crear. Esto ha venido a llamarse Web 2.0

Toda esta evolución desde unos pocos documentos estáticos hasta la infinidad de páginas dinámicas que tenemos ahora ha surgido según las necesidades de los usuarios de la web, y enfocados a estos.

Debido a esto las páginas web están pensadas para ser entendidas por humanos, y no es fácil hacer una clasificación automática de los contenidos que se encuentran en la red. Las búsquedas de información se hacen difíciles y normalmente un usuario tiene que mirar y descartar varios documentos hasta que encuentra lo que estaba buscando. A veces la información permanece oculta tras multitud de resultados de búsqueda sólo parcialmente correctos.

La Web Semántica es un intento de traer orden a este caos. Si se organiza la información en la web de forma que pueda ser entendida por un ordenador, se puede mejorar la efectividad de las búsquedas. Si además esta información se relaciona entre si podemos realizar búsquedas basadas en la semántica de la información buscada o incluso crear agentes que realicen gestiones complejas de forma automática.

2.2. Ontologías

El término ontología es el nombre de un campo de la filosofía que estudia la naturaleza de la existencia, identificando los tipos de cosas que existen y cómo se describen.

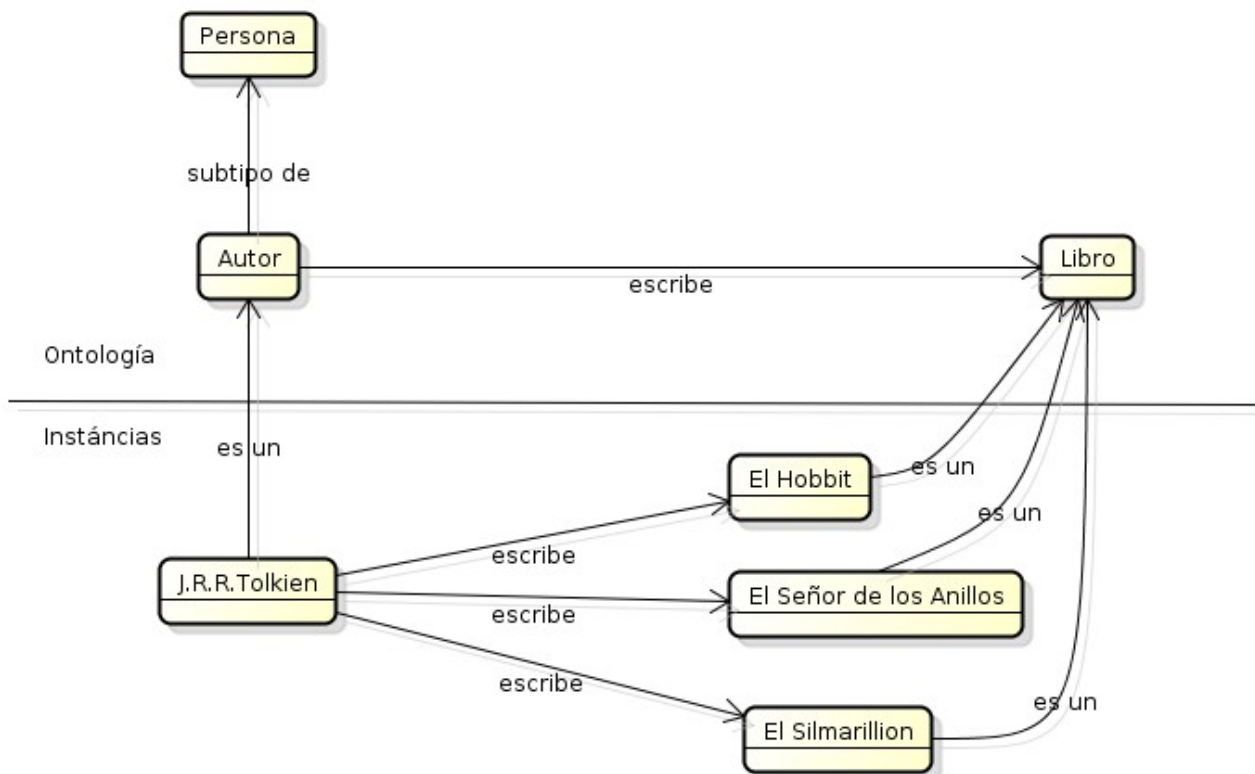
En el ámbito de la informática, se da un uso diferente. En vez de "ontología" nos referimos a "una ontología" que describe formalmente un dominio de interés, dando una lista finita de términos y su relación entre ellos. Estos términos muestran conceptos importantes (clases de objetos) del dominio. Por ejemplo, en el marco de la literatura se describen libros, autores, géneros literarios y otros conceptos.

Las relaciones normalmente muestran jerarquías entre los términos, indicando por ejemplo que la fantasía épica es un tipo de novela.

A parte de estas relaciones entre los términos, una ontología también muestra las propiedades que definen los términos. Por ejemplo, todo libro tiene un título, un autor tiene nombre, un libro ha sido escrito por uno o más autores...

Una ontología nos da un conocimiento formal sobre un dominio, y en la Red, una ontología nos da un conocimiento compartido sobre un dominio, permitiendo describir los elementos de un dominio de una forma común.

En la figura siguiente podemos ver un ejemplo simple de ontología referida a autores y libros. La parte superior de la figura muestra la ontología, donde se definen los tipos de elementos que tenemos en nuestro dominio (autores, personas y libros) y cómo se relacionan entre ellos (un autor es un tipo de persona, un autor escribe libros). La parte inferior de la figura muestra un uso de esta ontología definiendo un autor y tres de sus libros escritos.



2.3. XML y XML Schema

El primer paso para que la información sea entendible por un computador es estructurarla.

En HTML encontramos la información de un modo literario. Una página sobre un libro contendrá la información sobre quién es su autor mezclada con otra información redactada (por ejemplo "En <El Hobbit> J.R.R.Tolkien nos cuenta..."). Aunque a una persona le pueda resultar fácil intuir y extraer el nombre del autor a partir del contexto, a una máquina le resulta muy difícil.

XML es un lenguaje de marcado que permite la estructuración de la información en base a etiquetas (conocidas como tags) que indican cual es la información que contiene. La información anterior en XML podría ser:

```
<libro>
  <título>El Hobbit</título>
  <autor>J.R.R.Tolkien</autor>
</libro>
```

XML-Schema es un lenguaje que permite definir cual es la estructura que debe tener un determinado documento XML. De este modo, pueden definirse esquemas comunes para facilitar la interoperabilidad entre distintos documentos.

Podría pensarse que con XML se favorece la comprensión del documento por parte de las máquinas a costa de empeorar la legibilidad. XSLT son documentos que pueden

adjuntarse a documentos XML y procesarlos, transformando su contenido a otro más fácilmente entendible por las personas. Esto tiene la ventaja adicional que de un mismo XML pueden surgir documentos diferentes (HTML,PDF) simplemente cambiando el documento de transformación XSLT que lo procesa.

XML es un metalenguaje universal para definir marcas de información. Provee un entorno uniforme y una paleta de herramientas como procesadores (parsers), para permitir el intercambio de datos entre aplicaciones.

2.4. RDF y RDF-Schema

Las sentencias en RDF

Hemos visto que XML permite el intercambio de información, pero no dice nada sobre el significado (la semántica) de esta información. Es responsabilidad de cada aplicación el interpretar los tags de un documento XML.

Por ejemplo, supongamos que tenemos un XML que muestra una lista de personas, convenientemente etiquetadas como <persona>. Prestando atención al XML de libros que hemos mostrado antes. ¿Es J.R.R.Tolkien una persona?

RDF es un modelo de datos que permite definir información de manera semántica. Esta información se organiza en forma de tripletes del tipo sujeto-predicado-objeto, llamados sentencias (statements).

Utilizando el anterior ejemplo del libro, en RDF tendríamos dos sentencias

```
"El hobbit" es un libro
"El hobbit" tiene como autor a J.R.R.Tolkien
```

Si comparamos esta forma de definir la información con el modo en XML, vemos que ,mientras que en XML tenemos un dato "título" con valor "El hobbit" y un dato "autor" con valor J.R.R.Tolkien, en RDF estamos relacionando información añadiendo semántica.

Amplieemos el ejemplo mediante una nueva sentencia

```
"El señor de los anillos" tiene como autor a J.R.R.Tolkien
```

Podemos hacernos la pregunta "¿Cuáles son los libros que tienen como autor a J.R.R.Tolkien?". A partir de las sentencias anteriores, vemos que esos libros son "El hobbit" y "El señor de los anillos".

Hemos visto a grandes rasgos cómo RDF nos permite definir información, y también cómo se puede buscar información en esta estructura, pero ¿Es J.R.R.Tolkien una persona?. Antes de responder a esta pregunta debemos ver cómo funciona la siguiente capa, RDF-Schema

RDF-Schema como herramienta para la creación de ontologías

RDF-Schema (RDFS) es un lenguaje que permite definir la terminología que va a ser empleada en un determinado contexto. RDF se encarga de utilizar dicha terminología para establecer relaciones creando así una ontología.

RDFS permite añadir información semántica sobre cómo se relacionan entre sí las clases de nuestra ontología, estableciendo una jerarquía entre ellas.

Clases, propiedades e instancias

La base para la definición de ontologías mediante RDF se encuentra en la definición de clases, propiedades e instancias.

Una clase indica una categoría característica dentro de nuestra ontología. En una ontología literaria "Libro" sería una clase, así como también "Autor". Una clase contendrá instancias, que son los elementos de esa clase. La clase se relaciona con la instancia respondiendo a la pregunta "¿Qué es esta instancia?". De este modo "J.R.R.Tolkien" es un "Autor", mientras que "El Hobbit" es un "Libro".

En RDFS, las clases permiten crear jerarquías entre ellas (taxonomías), respondiendo a la sentencia "es un subtipo de". Siguiendo con el ejemplo "Autor" es un tipo de "Persona". Esto indica que todo autor es también una persona y comparte todos los atributos que tienen las personas.

Las propiedades permiten definir los atributos de una clase, estableciendo relaciones entre clases o entre clases y valores. En nuestro caso un "Libro" "tiene un" "Autor" (o más de uno), o un "Libro" "tiene como nombre" "un texto"

Así podemos ampliar el ejemplo con la siguiente sentencia indicando la relación comentada.

```
Autor es un subtipo de Persona
```

Dominios y rangos

Si nos fijamos en una sentencia de RDF , formada por sujeto-predicado-objeto y cuyo objeto sea una propiedad, podemos fijarnos que no hay información sobre qué debe ser el sujeto o el objeto.

RDFS permite definir esta información. El sujeto es el dominio de la propiedad, mientras que el objeto es el rango. Ampliaremos nuestro ejemplo con las siguientes sentencias en RDFS

```
la propiedad "tiene como autor" tiene como dominio un Libro
la propiedad "tiene como autor" tiene como rango un Autor
```

Así, de la sentencia del ejemplo

```
"El hobbit" tiene como autor a J.R.R.Tolkien
```

podemos saber que "El hobbit" es un Libro (por el dominio de la propiedad) y que J.R.R.Tolkien es un Autor. Este proceso de descubrir información a partir de las definiciones y sentencias disponibles es lo que se entiende como Inferencia.

Inferencia de nueva información a partir de información semántica

Podemos entender la inferencia como la creación de nuevas sentencias a partir de las sentencias existentes. Desde el punto de vista de la información, es la creación de nuevos conocimientos a partir de la información ya conocida y es donde reside la

verdadera potencia de la tecnología de web semántica.

A partir de la aplicación de reglas simples en las sentencias que tenemos, podemos inferir nuevas sentencias.

En este punto estamos preparados para utilizar la inferencia para responder a la pregunta que hacíamos al comenzar el ejemplo. ¿Es J.R.R.Tolkien una persona? Partimos de las sentencias

```
"El hobbit" tiene como autor a J.R.R.Tolkien
La propiedad "tiene como autor" tiene como rango Autor
```

para inferir la sentencia

```
J.R.R.Tolkien es un Autor
```

y dado que sabemos que

```
Autor es un subtipo de Persona
```

podemos inferir que

```
J.R.R.Tolkien es una Persona
```

La inferencia es una herramienta poderosa al permitir la creación de nuevo conocimiento a partir del conocimiento existente y las reglas definidas, pero debe ser observado con cuidado ya que una regla mal utilizada puede dar lugar a sentencias incorrectas.

2.5. OWL

La expresividad de RDF/RDF-Schema es limitada. Hemos visto como RDF permite definir predicados binarios ("J.R.R.Tolkien" es un "Autor"), mientras que RDF-Schema amplía la expresividad de RDF permitiendo la definición de jerarquías de clases de elementos y propiedades y sus relaciones.

El grupo de trabajo de ontologías web del W3C identificó un número de casos de uso para la Web Semántica que requieren mayor expresividad de la que provee RDF/RDF-Schema. El trabajo de varios grupos de investigación en Europa y Estados Unidos, que habían identificado estas necesidades y trabajado en una solución, se unió como punto de partida para OWL(Ontology Web Language), que está enfocado a ser el lenguaje estándar y ampliamente aceptado para la Web Semántica.

OWL ofrece mayor expresividad de la que tiene RDF/RDF-Schema, permitiendo crear ontologías más ricas.

Entre las características que aporta OWL encontramos la posibilidad de definir restricciones de cardinalidad en propiedades (una persona tiene exactamente dos padres) o características especiales de propiedades como la propiedad transitiva o la inversa.

Para entender mejor la expresividad de OWL, volveremos a la pregunta "¿Cuáles son los libros que tienen como autor a J.R.R.Tolkien?". Si volvemos a formular la pregunta del modo "¿Qué libros ha escrito J.R.R.Tolkien?" vemos que nos falta definir la propiedad "ha escrito" para poder responderla.

Aunque podríamos definir esta propiedad y ampliar el ejemplo con las sentencias

```
J.R.R.Tolkien ha escrito "El hobbit"  
J.R.R.Tolkien ha escrito "El señor de los anillos"
```

es algo engorroso teniendo en cuenta que ya teníamos que

```
"El hobbit" tiene como autor a J.R.R.Tolkien
```

La expresividad de OWL permite superar este escollo. Mediante la definición de la propiedad "ha escrito" como inversa de la propiedad "tiene como autor"

```
"ha escrito" es una propiedad inversa de "tiene como autor"
```

podemos inferir las sentencias anteriores y también la sentencia

```
"tiene como autor" es una propiedad inversa de "ha escrito"
```

que permitiría que al ampliar el ejemplo con

```
J.R.R.Tolkien ha escrito "El Silmarillion"
```

dando lugar por inferencia a la sentencia

```
"El Silmarillion" tiene como autor a J.R.R.Tolkien
```

3. Recursos de interés turístico

Siendo el objetivo de este trabajo la creación de una ontología turística geográfica, el primer paso es analizar nuestro dominio, saber cual es nuestro punto de partida y qué queremos obtener, y estudiar cuales son los elementos que nos encontraremos y qué relaciones hay entre ellos.

3.1. Buscando información sobre puntos de interés turísticos

En Internet existen una infinidad de recursos sobre puntos de interés turístico. Simplemente buscando "Catedral de Palma" en Google, surgen más de un millón de resultados para esta búsqueda, cada uno de ellos variando ampliamente en formato y tipos de información mostrada.

En Wikipedia podemos encontrar su historia y sus características en formato wiki. En Google vídeos encontramos fotos y vídeos de diferentes aspectos (ceremonias, arquitectura, eventos), y en la web institucional de la Catedral podemos encontrar sus horarios y actividades.

3.2. El problema de la información fragmentada

La Catedral de Palma es un punto de interés turístico muy importante y, como tal, dispone de un abanico de recursos en internet bien clasificados y diferenciados, pero cuando buscamos información sobre otros puntos de interés vemos que la calidad de la información es menor y mas dispersa, y que cuesta bastante localizar la información específica que necesitamos. En muchos casos tenemos que entrar en varios enlaces para localizar la información buscada.

Esta búsqueda se agrava cuando en lugar de centrar nuestra búsqueda en un punto de interés específico, lo que intentamos localizar es una lista de ellos que cumplan determinados criterios. En estos casos difícil encontrar la información de una forma no segmentada según las páginas.

3.3. Una ontología turística

Como punto de partida tenemos a nuestra disposición un amplio catálogo de recursos digitales sobre puntos de interés turístico, pero estos recursos no se encuentran clasificados.

Las ontologías han demostrado ser una tecnología clave para un acceso efectivo a la información, ya que permiten superar algunos de los problemas de la búsqueda basada en textos, relacionando y agrupando elementos según propiedades complejas. Por ello el camino a seguir es la creación de una ontología turística adecuada a nuestro dominio, que clasifique tanto los recursos digitales de los que disponemos, como los diferentes puntos de interés sobre los que trata cada uno de estos recursos.

3.4. Objetivos de la ontología

Antes de comenzar a diseñar nuestra ontología es necesario estudiar qué se pretende conseguir con ella. Dado que queremos mejorar la clasificación de los recursos digitales, estudiaremos nuestros requerimientos en base a unas supuestas preguntas que haría un usuario de nuestra ontología. Si sabemos qué se pregunta, sabremos cómo responder a esas preguntas.

- ¿Qué información hay sobre edificios religiosos en Palma de Mallorca?
- ¿Dónde están situados los centros comerciales?
- ¿Qué información de carácter comercial tenemos sobre el museo Es Baluard?
- ¿Cuáles son los vídeos disponibles sobre la Playa del Arenal?
- ¿Disponemos de la ficha de datos sobre la Catedral de Palma?

De estas preguntas comenzamos a desgranar los diferentes componentes que utilizaremos. Tenemos lugares como "Palma de Mallorca", puntos de interés como "Es Baluard" o la "Playa del Arenal", y recursos digitales como "Vídeos". Los puntos de interés tienen una clasificación como "Museo" o "Edificio religioso", mientras que los recursos digitales muestran una información como "de carácter comercial" o "ficha de datos"

3.5. Recursos digitales de interés turístico

El enfoque que se ha dado a este proyecto, centra el origen de la ontología en los recursos digitales. Esto es así porque es el material base del que disponemos.

A grandes rasgos, podemos definir los recursos digitales como documentos digitales accesibles a través de Internet (y por lo tanto identificados con una URL), y que pueden estar situados en diferentes servidores.

El dónde se encuentra un determinado recurso digital no debe preocuparnos ya que nuestra tarea es asegurarnos que el recurso digital queda bien referenciado e identificado, pudiendo indicarse tanto el qué es (PDF, imagen, vídeo, etc...) como el qué tipo de información muestra (descriptiva, comercial, divulgativa, etc...)

Formato de un recurso digital

La primera consideración sobre nuestros recursos digitales es su formato. El formato responde a la pregunta "¿Qué es?"

Según su formato, podemos encontrar vídeos, imágenes, audio, textos, PDFs, Wikis, HTML y otros tipos de multimedia. Cada uno de estos tipos de documento viene identificado en nuestra ontología mediante una jerarquía de clases, y cada instancia que creemos estará identificada por la URL del recurso digital.

Siguiendo esta dirección, cualquier documento digital disponible en Internet puede ser referenciado con este modelo, y pese a que nosotros nos centraremos en la clasificación de unos pocos recursos digitales a fin de mostrar el funcionamiento de la ontología, un proyecto de gran alcance podría rastrear recursos en la red mediante Crawlers y clasificarlos automáticamente.

3.6. Puntos de interés turístico

Un recurso es sólo un archivo digital, un contenedor. Si nos fijamos en el contenido de este archivo encontramos la información de carácter turístico. Esta información puede ser sobre distintos temas como estadísticas de visitantes, trámites, eventos, lugares. Es este último tema en el que nos centramos para este proyecto. En particular nos centraremos en recursos digitales que muestren información sobre puntos de interés turístico.

Un punto de interés turístico es un lugar físico con interés para una visita turística pertenecientes diversas tipologías. Podemos encontrar bares, playas, comercios, edificios religiosos, parques naturales, miradores, etc...

Determinados recursos digitales pueden referirse a elementos como actividades y eventos con interés turístico, o incluso a cuestiones más abstractas como gastronomía y productos gastronómicos. Pese a ser elementos con un gran interés para el turismo, el alcance de este proyecto hace que no podamos extendernos hacia la clasificación de este tipo de elementos, por lo cual quedan fuera de nuestro dominio. La ontología no tratará temas abstractos como gastronomía, historia, etc..., centrándonos en el ámbito de los puntos de interés como un lugar físico.

Es importante tener definida una taxonomía que organice los diferentes puntos de interés, creando una jerarquía de clases capaz de contener las instancias que crearemos y que cumpla los objetivos facilitando la búsqueda de las mismas.

Como lugar físico que es, un punto de interés (PdI) puede estar localizado por sus coordenadas geográficas. Las propiedades de latitud y longitud permitirán georreferenciar los PdI para que puedan ser mostrados en un mapa.

Un determinado PdI puede pertenecer a una clasificación determinada, pero tener usos diferentes. Por ejemplo el Gran Hotel de Palma es un edificio emblemático del modernismo que es usado como galería de arte, centro de reuniones y como bar-restaurante. Una propiedad "usadoComo" permitirá definir estas características especiales.

Una visión geográfica

Todo PdI se encuentra en algún lugar. Además de la posibilidad de incluir las coordenadas geográficas de un PdI, la ontología debe definir los ámbitos geográficos necesarios para localizarlos según su zona.

Es necesario que exista una taxonomía adecuada a tal efecto, definiendo conceptos como país, provincia, municipio, zona turística, etc...

Los PdI pueden (y por conveniencia deben) estar localizados mediante el uso de estas estructuras y una propiedad "estaSituadoEn", definiendo así dónde se encuentran en términos de zona o ámbito geopolítico. Esta propiedad es de tipo transitivo, de forma que la inferencia permita definir la pertenencia a las zonas a partir de las relaciones entre estas. De este modo, un determinado PdI que esté situado en la ciudad de Palma de Mallorca, mediante inferencia, estará situado debido a esto en la Isla de Mallorca y en España.

Más allá de este uso, esta misma propiedad permite indicar que un determinado PdI

forma parte o se encuentra en otro PdI, permitiendo que elementos actúen como contenedores de otros elementos (por ejemplo un cine situado en un centro comercial)

Relación entre los puntos de interés y los recursos digitales

Mediante una propiedad en la ontología, podemos relacionar los recursos digitales y los puntos de interés, indicando sobre qué puntos de interés trata un determinado recurso digital.

Creando también la propiedad inversa, podremos indicar cuales son los recursos digitales que tratan un determinado punto de interés.

Tipos de información

Volviendo a los recursos digitales, se han clasificado los recursos digitales según su formato, respondiendo a la pregunta "¿Qué es?". También se han relacionado con los puntos de interés respondiendo a la pregunta "¿Sobre qué trata?". Ahora se añadirá una propiedad para responder a "¿Qué tipo de información contiene?"

Es diferente un PDF con información comercial de un PdI que uno que contenga una ficha técnica. Por esto esta propiedad nos permite una clasificación adicional de los recursos a fin de poder filtrar sólo los que tengan interés para el usuario.

Para este efecto se creará una clase "Tipo de Información" que contendrá instancias de los diferentes tipos de información. Una lista inicial de elementos de esta clase serían: Información descriptiva, ficha de datos, información comercial, mapa de localización.

4. Ontología turística geográfica

En el análisis hecho en el capítulo anterior se establecen las guías a seguir para la creación de nuestra ontología turística geográfica. Necesitamos una taxonomía de puntos de interés, una taxonomía geopolítica para definir las zonas, una taxonomía de recursos digitales en Internet y enlazarlo todo de modo que forme una unidad.

4.1. Recursos semánticos en Internet

El primer paso para la creación de nuestra ontología es buscar qué recursos hay disponibles en Internet. Si conseguimos encontrar una o más ontologías que nos permitan anotar la información que necesitamos, conseguiremos diversas ventajas.

En primer lugar estaremos ahorrando trabajo. La ontología puede importarse en la nuestra y usar sus definiciones como si fuesen propias, evitando la tarea de clasificación y creación de una taxonomía, aunque sustituyéndola por la necesaria tarea de revisión.

La ontología importada puede extenderse. Si echamos en falta clases, o propiedades, o necesitamos dividir o renombrar términos, podemos hacerlo con un esfuerzo mínimo, consiguiendo así que se adapte completamente a nuestras necesidades.

Una ontología creada por otros que trate el mismo tema que nuestro proyecto nos permite mirar el tema objetivo desde un punto de vista distinto, y a veces nos hace pensar en entidades, términos y situaciones en las que no nos habríamos fijado de no tener este punto de vista externo.

Pero la ventaja más prometedora que conseguimos es la interoperabilidad. Podemos beneficiarnos de las instancias construidas para la ontología importada, mientras que otros pueden beneficiarse de gran parte de las instancias que construyamos en la nuestra. El uso de una ontología importada hace fuerte nuestra ontología.

Como última anotación, cuando comentábamos brevemente la historia de la WWW, vimos que la Web Semántica nació como un intento de traer orden al caos de formatos diferentes en Internet. La Web Semántica fue diseñada para permitir y fomentar la interoperabilidad entre diferentes fuentes de información, así que seguir este principio es, ante todo, una buena idea.

Una ontología en varios documentos

En la búsqueda en Internet de ontologías que puedan satisfacer nuestras necesidades, se ha localizado una que cumple casi a la perfección con los requerimientos que teníamos en los Puntos de Interés y lugares. En el apartado siguiente explicamos la estructura de esta ontología, pero esta ontología no cubre completamente nuestras necesidades ya que, además de los ajustes necesarios, tenemos que modelar una taxonomía de recursos digitales.

Se ha decidido crear una ontología separada para los recursos digitales. Esto es debido a que puede ser de interés para otras personas este modelo y el hecho de mantenerlo separado hace que si otro desarrollador desea crear una ontología que clasifique recursos digitales que no tengan un carácter turístico, no tenga que cargar con la terminología turística que a nosotros nos es útil pero para otros puede ser un estorbo.

De este modo, nuestra ontología está compuesta por tres documentos. Uno que modela puntos de interés y lugares, otro que modela recursos digitales, y un tercero que importa los otros dos y hace de enlace entre ellos, extendiendo y modificando las partes de la ontología reutilizada que no se adaptaban a nuestras necesidades.

En los apartados siguientes profundizamos en la estructura de cada uno de estos documentos.

4.2. Puntos de Interés: Space.owl

Esta ontología, que fue encontrada en Internet mediante el buscador de ontologías Swoogle, modela puntos de interés y lugares. Aunque se incluye una copia de el archivo OWL en este proyecto, su origen y las referencias a la hora de importarla a nuestra ontología se hacen mediante la URL de origen

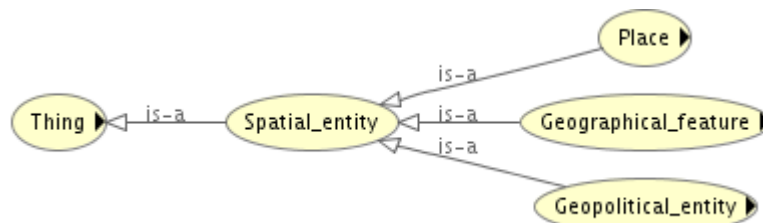
<http://212.119.9.180/Ontologies/0.3/space.owl>

Es una ontología que modela diferentes conceptos espaciales y sus descripciones, que coinciden con los requerimientos que teníamos para los Puntos de Interés y para la taxonomía geopolítica.

En este apartado estudiamos su estructura. Dado que esta ontología ha sido creada en inglés, nos referiremos a los términos en inglés, junto con la traducción (literal o conceptual) entre paréntesis donde sea necesario.

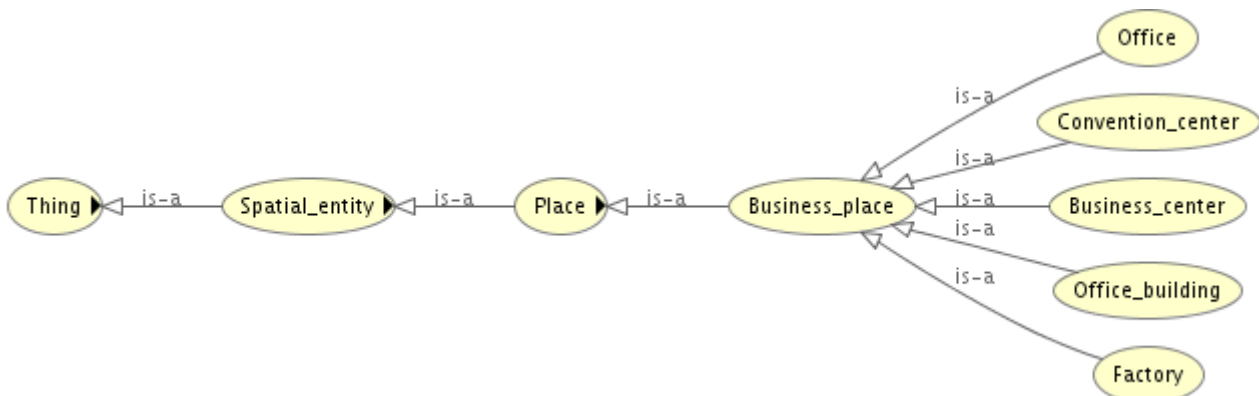
Estructura general de las clases

Existe una clase raíz que modela el concepto de entidad espacial (*Spatial_entity*) de la cual extienden las clases *Place*, *Geographical_feature* y *Geopolitical_entity*, que se refieren respectivamente a espacios contruidos por el ser humano (museos, catedrales...), espacios naturales (playas, montañas...) y divisiones geopolíticas (países, ciudades, ...)



Espacios contruidos por el ser humano: Place

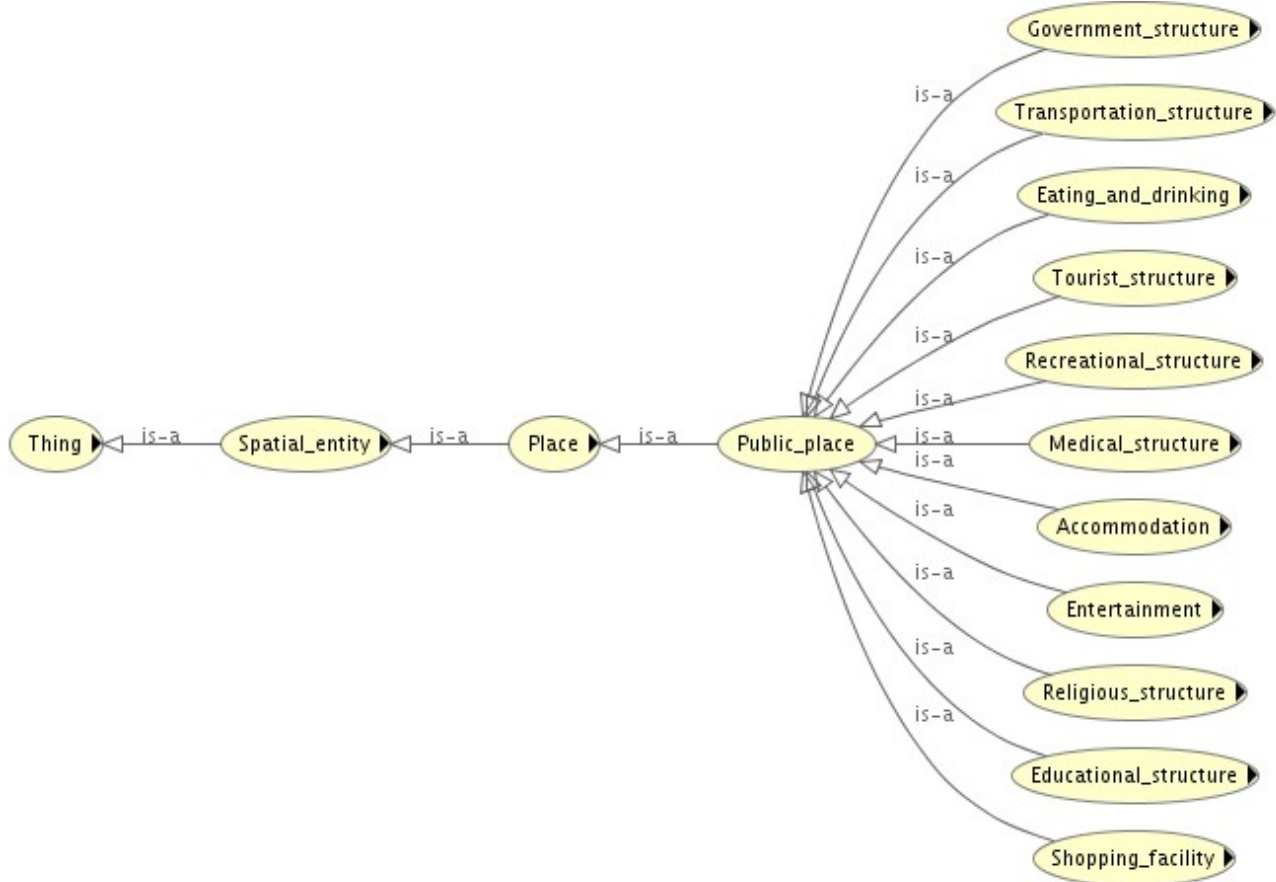
Bajo esta clase se encuentran los diferentes tipos de punto de interés que ha construido el ser humano. Se dividen en tres tipos *Business_place* (espacios de negocios), *Private_place* (espacio privado) y *Public_place* (espacio público). Para nuestra ontología en principio no tendrá utilidad la clase *Private_place* y sus descendientes ya que no se prevé ningún punto de interés de este tipo.



Los espacios de negocios tienen una utilidad limitada, aún así los centros de convenciones (*Convention_center*) y las fábricas (*Factory*) tienen un gran potencial para nuestro proyecto.

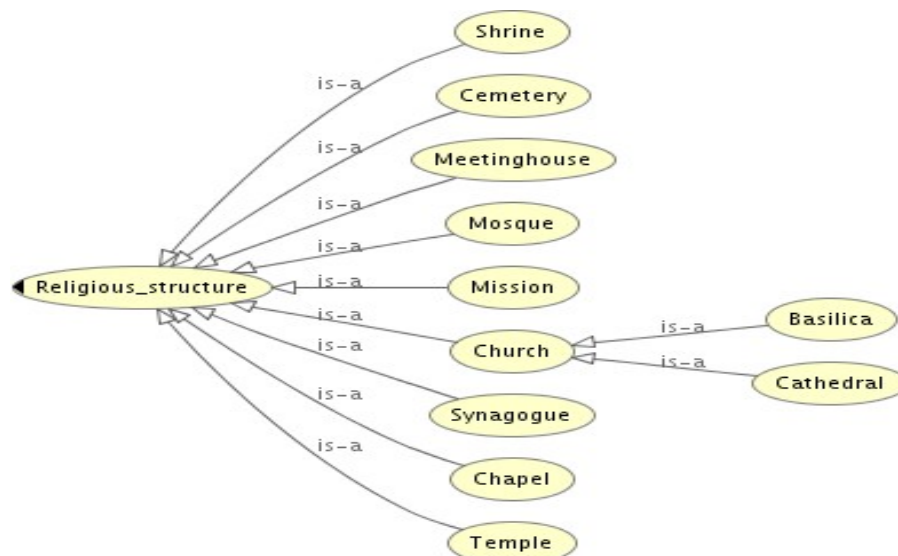
Los espacios públicos se ajustan como un guante al proyecto. Es una taxonomía extensa que cumple perfectamente con lo que necesitamos para definir nuestros puntos de interés. Desde el área directamente turística con museos y monumentos, hasta el área de negocios con tiendas y grandes almacenes, y pasando por otras áreas

claves como movilidad y transporte , restauración, alojamiento y entretenimiento.



En la figura que se incluye a continuación se muestran las áreas de las que se componen los espacios públicos.

Cada una de estas áreas modela diversas clases relacionadas con el área. A modo de ejemplo, en la figura siguiente se muestra la estructura que desciende de Religious_structure (Estructura religiosa) ya que esta ha sido fuente de ejemplos a la hora de definir los requerimientos.



Espacios naturales: Geographical_feature

Descienden de esta clase todos los puntos de interés que tienen un origen natural o con una relación directa con la parte natural.

Como puede observarse en la página siguiente, no existen los conceptos que tienen relación con divisiones políticas de espacios naturales, como son los parques naturales, zonas de caza, etc... En este proyecto se ha optado también por no incluirlas para ajustarnos lo más posible al original. Una clasificación de este tipo se escapa un poco del carácter de puntos de interés que define nuestro proyecto. A fin de cuentas un bosque que es un parque natural sigue siendo un bosque.



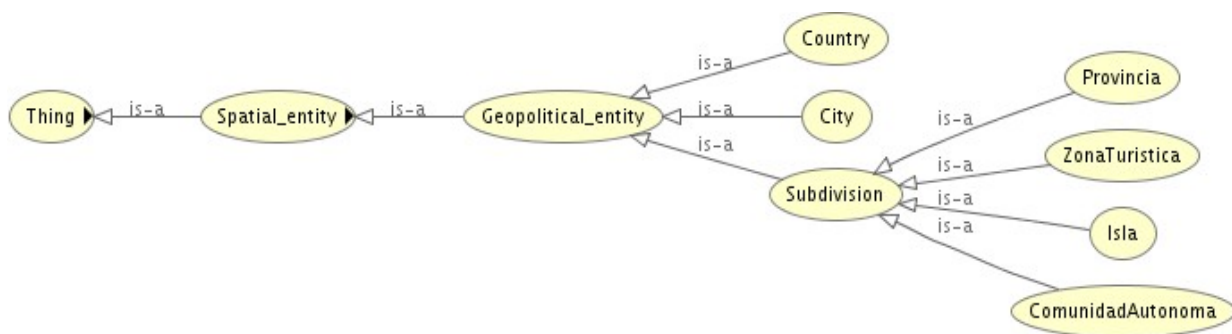
Estructura política y zonal: Geopolitical_entity

Como comentábamos al inicio de este apartado, la ontología soluciona dos de las taxonomías que habíamos detectado en el análisis; los puntos de interés y los lugares.

Esta clase se refiere a esta última taxonomía.

La taxonomía únicamente incluía Country (país), city (ciudad) y Subdivision (subdivisión). Esto resulta algo pobre para la estructuración que necesitamos en este punto, por lo que podemos observar que se han creado clases descendientes de subdivisión que nos permiten el grado de detalle necesario.

Esta extensión a la ontología spaces.owl se ha realizado en el documento RecursosTuristicos.owl que, como veremos en un apartado posterior, extiende esta ontología y la enlaza con la ontología de recursos digitales.



Definiciones exclusivas y no exclusivas

La ontología sólo define como clases disjuntas a Geopolitical_entity con las otras dos superclases hermanas. Esto indica que un mismo elemento no puede ser a la vez una división geopolítica y un espacio natural o construido por el ser humano.

Viendo este hecho desde el punto de vista contrario, la ontología permite que un mismo elemento pertenezca a más de una clase. Rescatemos el ejemplo del Gran Hotel de Palma, y hagamos que este elemento pertenezca a las clases Museum(museo), Bar, Restaurant (restaurante) y Convention_center (centro de reuniones) para clasificarlo con precisión.

4.3. Recursos digitales: RecursosDigitales.owl

Esta ontología ha sido creada expresamente para este proyecto. Además de la copia de el archivo OWL que se incluye en el proyecto, esta ontología puede encontrarse en la URL

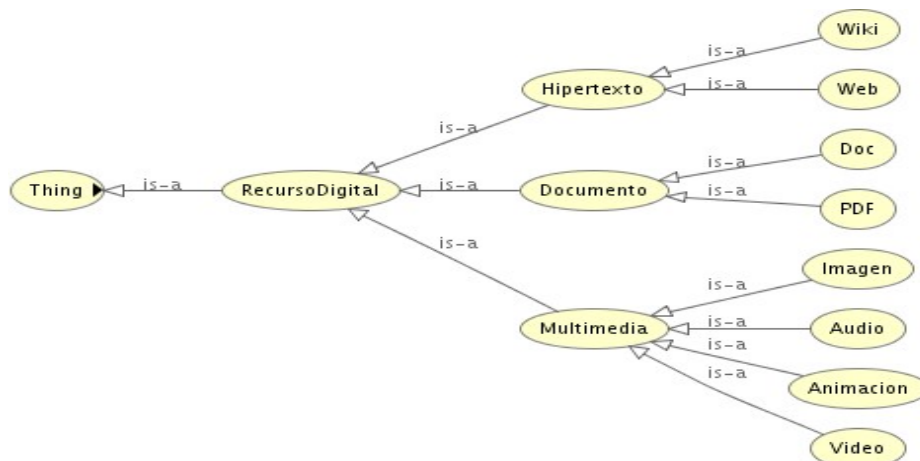
<http://www.aybkamen.com/arioch/ontology/RecursosDigitales.owl>

El dominio de esta ontología son los diferentes tipos de documentos que podemos encontrar en Internet y su tipo de contenido. Esta ontología define hipertextos, multimédias y documentos como los tipos de recursos que pueblan Internet, estableciendo algunos subtipos muy comunes como subtipos de estos.

Además de clasificar los tipos de recursos digitales, esta ontología permite definir

qué tipo de contenido tiene el recurso digital. Mediante el uso de la clase TipoInformacion, un usuario puede definir cuales son los tipos de información disponibles en su dominio (por ejemplo información comercial, datos técnicos, etc.). La propiedad tieneInformacion permite indicar cuál es el contenido de un recurso digital. De este modo una determinada Wiki puede tener información de datos técnicos y un vídeo puede tener información comercial.

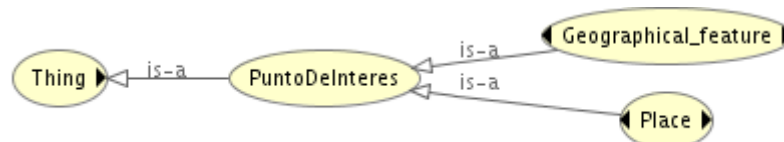
Se ha diseñado esta ontología pensando en que pueda ser reutilizada. Por ello sólo incluye cuestiones referentes a recursos digitales, sin incluir ninguna referencia a puntos de interés o cuestiones turísticas. Esto es así para que un futuro usuario esté libre de conceptos que no necesita, pudiéndola utilizar como un componente para la creación de una ontología que no tenga relación con el turismo.



4.4. Ontología turística geográfica: RecursosTurisiticos.owl

La creación de esta ontología, que puede ser considerada como el tronco de este proyecto, tiene como objetivo el juntar las dos ontologías anteriores y extenderlas, de modo que estén relacionadas entre sí. De este modo, juntando una ontología de ámbito turístico con una tecnología sobre recursos digitales, conseguimos una ontología sobre recursos digitales de interés turístico. Esta extensión de las ontologías base se hace de tres modos.

La primera extensión es crear una clase PuntoDeInteres y hacer que todos los tipos de punto de interés hereden de ella. De este modo, y amparados en la inferencia, podemos realizar una búsqueda sobre esta clase y obtener todos los puntos de interés de nuestro sistema. Sin esta clase, nuestros puntos de interés estarían divididos entre los que descienden Geographical_feature y los que descienden de Place, obligándonos a hacer una búsqueda compuesta para obtener todos los PdI.



La segunda extensión viene de definir la propiedad transitiva estaSituadoEn. Esta propiedad permite jerarquizar nuestras instancias, estableciendo relaciones de contenido entre ellas. Esta estructura nos permitirá realizar las consultas por zona

geográfica, sin importarnos a qué grado de detalle nos referimos. Por ejemplo, una consulta sobre los PdI situados en una provincia nos devolverá como resultado un cine situado en un centro comercial de una ciudad de un municipio de dicha provincia. El definir esta propiedad como transitiva es lo que activa la inferencia. Sólo hemos de preocuparnos de definir que ese cine está situado en el centro comercial, para que la magia de la inferencia haga el resto.

La tercera extensión relaciona puntos de interés con recursos digitales mediante la propiedad muestra. Un determinado recurso digital muestra uno o varios puntos de interés.

Con esta ontología como nexo de unión entre las dos anteriores completamos nuestro objetivo. El resultado final es una ontología que nos permite modelar recursos digitales, lugares geográficos y puntos de interés, y relacionarlos entre ellos, creando de este modo una herramienta para la creación de sistemas de información de carácter turístico geográfico.

5. Buscador de recursos turísticos

5.1. Requerimientos de la aplicación

El objetivo principal es construir una aplicación web que muestre en funcionamiento la ontología creada, permitiendo buscar y navegar entre los diferentes puntos de interés instanciados.

El usuario de la aplicación ha de poder localizar los puntos de interés según su tipología, pero el usuario no ha de saber necesariamente a qué tipología exacta pertenece un determinado PdI, de modo que puede pedir qué instancias pertenecen a una clase, y el sistema ha de devolver en los resultados de la búsqueda los puntos de interés pertenecientes a esa clase o a una subclase de esta en la jerarquía. Por ejemplo, pedir información sobre Edificios Religiosos ha de devolver cementerios, catedrales, iglesias, ...

Otro tipo de búsqueda disponible es por localización. Al igual que en el caso anterior, cuando un usuario pida información sobre un lugar, ha de devolverse la información de ese lugar y todos los lugares que forman parte de él.

El usuario puede buscar por nombre, escribiendo parte del nombre del punto de interés para filtrar la búsqueda.

Los tres tipos de búsqueda anteriores son combinables, de forma que puede realizarse una consulta sobre qué puntos de interés contienen cierto texto, pertenecen a una tipología y están en un determinado lugar.

Los resultados obtenidos muestran una lista ordenada con los PdI que cumplen los criterios de la búsqueda, así como un recordatorio de esos criterios.

A partir de los resultados obtenidos, el usuario puede acceder a la ficha de cada uno de los recursos encontrados. En esta ficha se muestra el nombre y la clasificación del recurso, se muestran los diferentes recursos digitales asociados a ese recurso y un mapa de localización que muestra dónde está situado el recurso.

5.2. Análisis de funcionalidades

A partir de los requerimientos indicados en el apartado anterior, podemos confeccionar la siguiente lista de funcionalidades.

1. La aplicación ha de acceder a la ontología creada.
2. La aplicación realiza las consultas sobre las instancias de ejemplo que se han creado.
3. El usuario puede buscar por tipología.
4. El usuario puede buscar por localización.
5. El usuario puede buscar por nombre.
6. El usuario puede buscar por tipo de recurso, indicando cuáles son los tipos de recurso requeridos.
7. La búsqueda por tipología localiza elementos con un nivel inferior en la jerarquía de clases.

8. La búsqueda por localización localiza elementos que están situados en la localización seleccionada o en las localizaciones que forman parte de esta.
9. El resultado de la búsqueda es una lista, ordenada por nombre, de los puntos de interés que cumplen con los criterios seleccionados por el usuario.
10. Cada elemento de la lista de resultados enlaza a la ficha de dicho elemento.
11. Junto al resultado de la búsqueda se muestran los criterios que han llevado a este resultado.
12. La ficha de un punto de interés muestra su nombre.
13. La ficha muestra las clases a las que pertenece el Pdl.
14. La ficha muestra las imágenes del Pdl.
15. La ficha muestra los vídeos del Pdl.
16. La ficha muestra las páginas web del Pdl.
17. La ficha muestra un mapa con la localización del Pdl.
18. La aplicación permite navegar por ella para volver a realizar otra consulta o para ver la lista de resultados de la última consulta.

Las funcionalidades 7 y 8 nos indican que probablemente va a ser necesario utilizar la inferencia para resolverlas. Veremos cómo en el diseño de la aplicación.

5.3. Diseño de la aplicación

Dado que estamos tratando con una aplicación web, el diseño de la misma ha de pensarse utilizando el concepto de páginas web. A partir de las funcionalidades de la aplicación, utilizaremos un diseño orientado al usuario para diseñar el interfaz de la aplicación, y a partir de este resolveremos las funcionalidades de sistema.

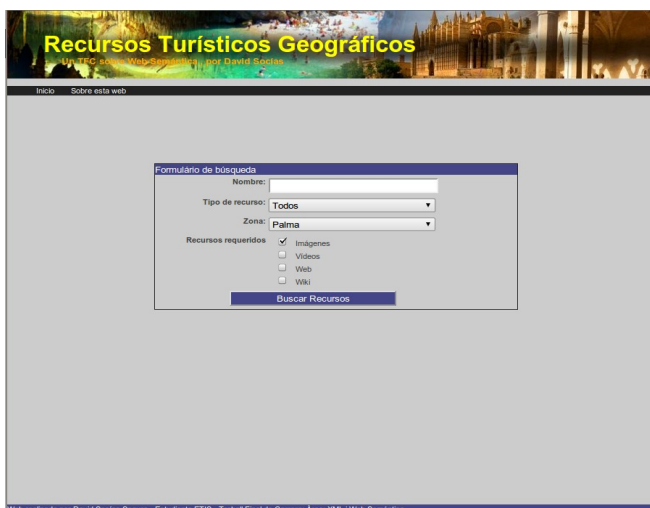
5.4. Interfaz de la aplicación

Podemos dividir nuestra aplicación en tres páginas.

La primera es la página de búsqueda. En ella se muestra un formulario donde se pide al usuario cuáles son los parámetros de la búsqueda. Este formulario contiene un cuadro de texto para indicar el filtro de nombre (funcionalidad 5), un selector que muestra la clasificación de los Pdl para que el usuario seleccione la clasificación requerida (funcionalidad 3) , un selector que muestra las diferentes localizaciones para realizar la búsqueda (funcionalidad 4) y un botón que lanza la búsqueda.

La página de resultados muestra la lista de los Pdl que cumplen los criterios de búsqueda (funcionalidad 9) así como un cuadro con qué criterios de búsqueda se han utilizado (funcionalidad 11). Cada uno de los elementos de esta lista de resultados enlaza a la ficha del elemento (funcionalidad 10)

La página de ficha muestra una ficha con los datos del Pdl, indicando su nombre, clasificación, imágenes, vídeos y páginas web enlazadas (funcionalidades 12 a 16). También muestra un mapa donde se indica la localización del Pdl (funcionalidad 17).



API para Web Semántica

De forma obvia, vamos a necesitar hacer uso de alguna API de Web Semántica para desarrollar nuestra aplicación. Aunque la decisión de cuál API en particular va a ser empleada corresponde a la arquitectura de la aplicación, a la hora de realizar el diseño necesitamos definir los requerimientos para esta API y suponer unas clases o estructuras que nos permitan trabajar.

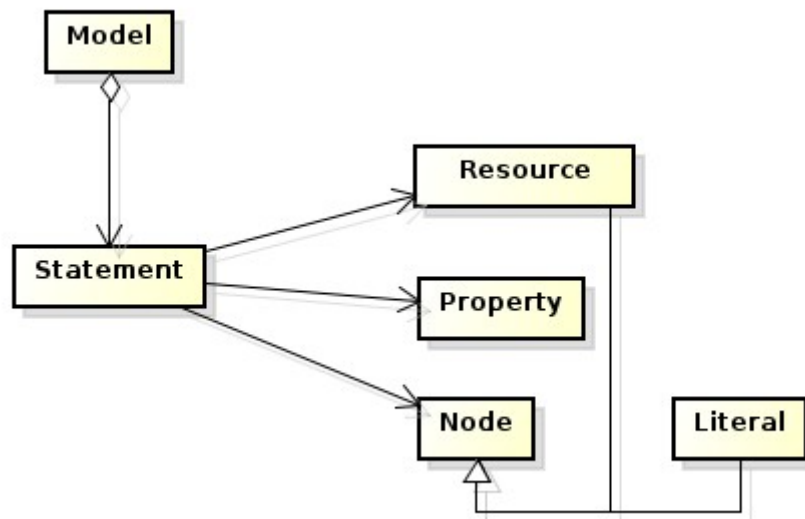
Los requerimientos que suponemos para la API son:

- Puede trabajar con RDF, RDFS y OWL.
- Permite utilizar un razonador para solucionar la inferencia.
- Permite navegar desde un elemento hasta otros a partir de sus sentencias y relaciones

Suponemos las siguientes clases:

- Model: Contiene todo un modelo de datos completo, con todas las sentencias del modelo
- Statement: Modela una sentencia de tipo Sujeto-predicado-objeto
- Resource: Indica un elemento de la ontología, así como un sujeto de una Statement.
- Property: Modela una propiedad de la ontología como `rdf:type` o `turistico:estaContenidoEn`
- Literal: Modela un valor literal utilizado como objeto de una Statement

El diagrama UML de las clases que suponemos es:



Debido a que el objeto de una sentencia puede ser un literal o un recurso, es necesaria una clase Node que actúe como superclase de Resource y Literal.

Clases descriptivas

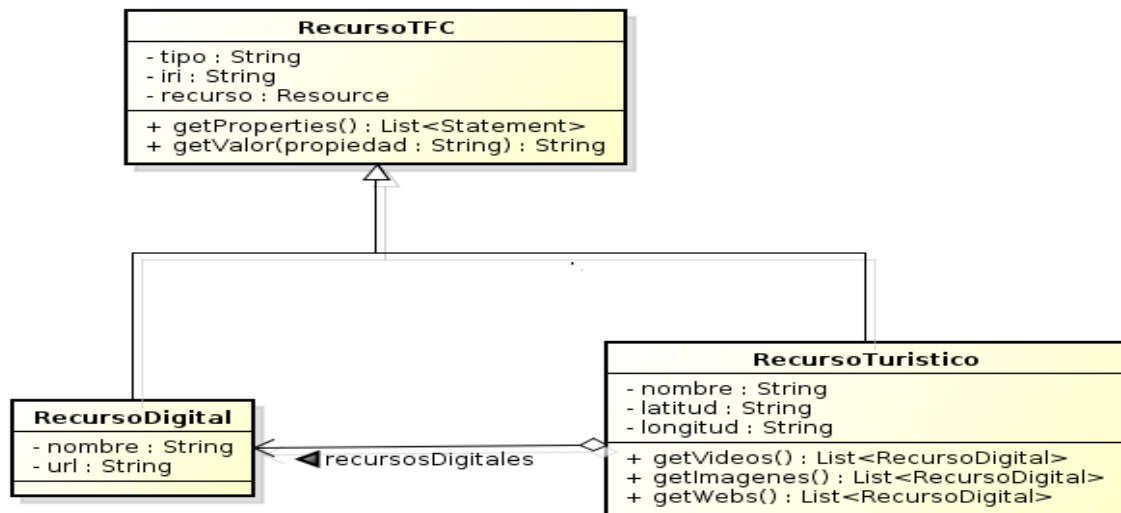
En nuestra aplicación vamos a tratar con puntos de interés y con recursos digitales que estarán definidos en el modelo semántico como Resources. Por ello es necesario transformar estos Resources para obtener unas clases más adecuadas a nuestras necesidades.

Hay que crear unas clases `RecursoDigital` y `RecursoTuristico` que modelen los elementos de nuestro dominio.

Como los valores de los atributos de estas clases provienen de un `Resource`, creamos también una clase `RecursoTFC` de la cual heredan las otras dos clases, que facilita la transformación.

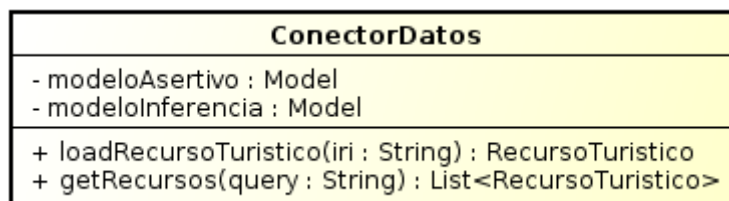
Estas clases sirven como adaptador, y serán utilizadas en la aplicación para referirnos a los recursos digitales y puntos de interés, evitando utilizar en la medida de lo posible la clase `Resource` de la API semántica.

La construcción de estas clases se hace a partir de su IRI (referencia) y de el modelo semántico.



Controlador de gestión

Para nuestra aplicación es necesario un controlador que permita gestionar el acceso al modelo semántico y que construya las clases descriptivas que solicitamos. La clase que se crea para este menester se llama `ConectorDatos` y su descripción es la siguiente:



Esta clase carga desde los documentos `.owl` el modelo asertivo (sin inferencia) y aplica un razonador básico para construir un modelo con inferencia.

El método `getRecursos` recibe como parámetro una consulta en el lenguaje de consultas SPARQL y devuelve la lista de `RecursoTuristico` que son resultados de esta

consulta.

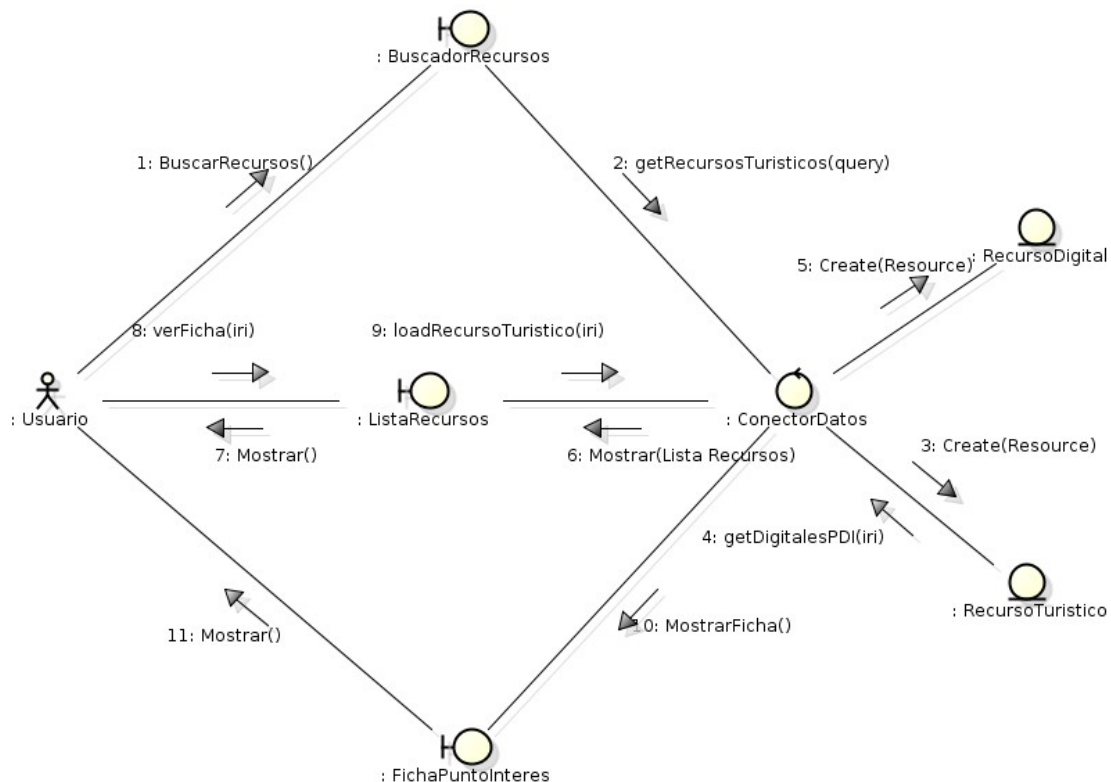
El método loadRecursoTuristico construye un RecursoTuristico completo a partir del identificador del recurso. Este método es utilizado en la consulta para construir los recursos.

Flujo de una consulta

El flujo principal para una consulta por parte del usuario es el siguiente:

1. El usuario rellena el formulario de consulta y pulsa del botón de buscar.
2. Se genera una sentencia de consulta en SPARQL con los datos de este formulario y se envía al método getResources de ConectorDatos.
3. La clase gestora resuelve la consulta sobre el modelo y crea un RecursoTuristico por cada resultado obtenido.
4. Cada RecursoTuristico le pide a ConectorDatos cuales son sus RecursosDigitales.
5. ConectorDatos recupera del modelo y crea los RecursosDigitales de cada RecursoTuristico.
6. Se guarda la lista de RecursoTuristico como resultado de la consulta.
7. Se muestra la página de lista de recursos al usuario.
8. El usuario indica que quiere ver la ficha de uno de los RecursosTurísticos.
9. Se carga el RecursoTuristico indicado.
10. Se guarda el RecursoTuristico como elemento seleccionado.
11. Se muestra la ficha de RecursoTuristico al usuario.

Este proceso puede verse resumido en el siguiente diagrama de colaboración.



5.5. Arquitectura del prototipo

Tecnología empleada

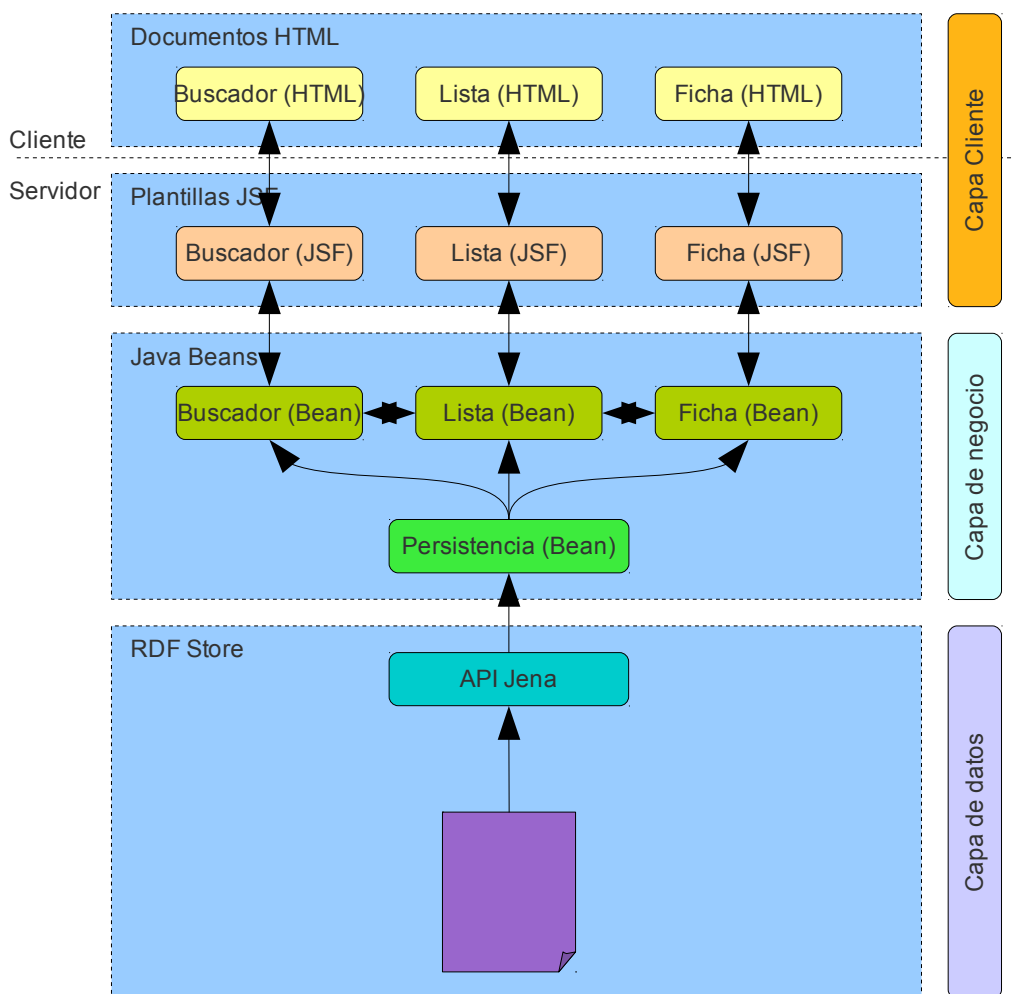
Para la creación del prototipo de buscador se usará Java-J2EE, utilizando Java Server Faces para la presentación de las páginas en el cliente y la comunicación con el servidor.

Para la gestión del modelo semántico se utilizara la API de web semántica Jena, que cumple los requerimientos indicados en el diseño de la aplicación.

Debido a que la aplicación simplemente realiza consultas y muestra resultado, y no es necesario modificar instancias de nuestra ontología desde la aplicación, no se ha utilizado ningún modelo de persistencia. La aplicación cargará los datos en memoria y trabajará con ellos sin alterarlos.

Estructura de capas

El siguiente diagrama muestra la estructura de la aplicación dividiéndola según las capas que la componen.



5.6. Instalación

Requerimientos de la aplicación

Para ser instalada, la aplicación necesita funcionar sobre un servidor de aplicaciones Java. Ha sido diseñada para funcionar sobre Tomcat 6.0 y, aunque en teoría puede ser ejecutada en otros tipos de servidor (Glassfish, Jboss), se recomienda el uso de este servidor de aplicaciones.

El servidor ha de estar correctamente instalado. Esto incluye el requerimiento de una máquina virtual Java. La aplicación ha sido desarrollada utilizando la plataforma JDK 1.6.

Los documentos OWL son cargados a través de su URL, por lo que es necesario que el sistema tenga acceso a la red para poder ejecutar la aplicación.

Requerimientos:

- Máquina virtual Java con soporte para la versión 1.6
- Servidor de aplicaciones Tomcat 6.x
- Acceso a Internet.

Opcionales:

- Aplicación Tomcat Manager para facilitar la instalación a partir de .WAR

Instalación a partir del .WAR

El documento RecursosTuristicos.war contiene el desplegable de la aplicación, listo para ser instalado en un servidor Tomcat.

Para instalarlo hay que copiar el documento en la carpeta de aplicaciones y el servidor se encargará de desplegarlo.

Alternativamente, si el servidor tiene instalada la aplicación Tomcat Manager, podemos acceder a ella para desplegar (Deploy) la aplicación.

Construir la aplicación a partir de las fuentes de la aplicación

El documento RecursosTuristicos.zip contiene toda la carpeta de proyecto de NetBeans para su instalación. Tras extraer los contenidos de este documento hay que importar el proyecto a NetBeans y corregir las referencias a carpetas erróneas. Es importante que el entorno NetBeans tenga soporte para proyectos Web y para JSF.

Otra posibilidad es utilizar los contenidos de este documento para crear un proyecto propio, sea en NetBeans, Eclipse, ANT o cualquier otro IDE. Para ello se describen a continuación las carpetas necesarias para construir un proyecto.

- **src/java** : Contiene los ficheros de la capa de negocio en forma de ficheros .java
- **web** : Contiene los ficheros de la capa cliente en forma de ficheros .xhtml

(JSF) , la carpeta CSS con la hoja de estilos y la carpeta WEB-INF que se indica a continuación

- **web/WEB-INF** : Contiene los ficheros de configuración del contexto (web.xml) y de Java Server Faces (faces-config.xml) así como la carpeta lib que se explica a continuación
- **web/WEB-INF/lib** : Contiene todas las librerías específicas de el proyecto, que son las de la API Jena y las de MyFaces.

Para crear el proyecto hay que crear un proyecto Java Server Faces, o un proyecto Web Dinámico que incluya JSF, e ir copiando los contenidos de las carpetas indicadas.

6. Conclusiones

En el desarrollo de este proyecto se ha construido una ontología simple y un prototipo básico de buscador de recursos. Aún en la simpleza de estos dos productos, se puede vislumbrar la potencia y elegancia de la tecnología de web semántica.

Durante una gran parte de mi vida laboral he estado ligado al ámbito turístico al trabajar para una consultora turística, desarrollando, como parte de mis proyectos, aplicaciones de soporte para centros de información turística. Estas aplicaciones eran simples colecciones de fichas, extremadamente rígidas, y sistemas de consulta sobre estas fichas. El tratamiento de los puntos de interés como fichas de datos crea bastantes problemas cuando nos encontramos con elementos que pertenecen a más de una clase, creando información duplicada; cuando nos encontramos relaciones entre elementos que no son comunes, pero que son necesarias en algunas fichas puntuales; cuando nos encontramos con una de las muchas excepciones que trae el modelar la realidad a las estructuras que habíamos creado. La solución de estos problemas la mayoría de veces trae un aumento importante de la complejidad del sistema, cargándolo de "parches" que sirven sólo para solucionar alguno de estos problemas de forma puntual.

6.1. Conclusiones sobre la ontología creada

Como puede verse en la ontología creada, la tecnología de web semántica supera estas dificultades con simpleza y elegancia. Debido a mi experiencia previa, mientras diseñaba la ontología mantenía la atención en cómo se superaban estos problemas, y el resultado va mucho más allá de lo esperado. No sólo se superan los problemas antes mencionados, sino que se incorporan otras ventajas.

La ontología creada es fácilmente extensible. Un usuario que necesite otras clasificaciones de PdI que no formen parte de la ontología original, puede crear un documento que defina estas clasificaciones y realizar una unión (merge) con la ontología actual, extendiendo su uso. Este mismo procedimiento puede usarse para incluir datos de ficha a los PdI (como indicar longitud de una playa o tipo de arena) o crear un puente de unión con otras ontologías externas y sus instancias, permitiendo enlazar con información de reservas de alojamientos, o carteleras de cines y espectáculos.

Al haber creado la ontología en trozos diferenciados se favorece el reuso de esta para proyectos que no tengan que ver con este proyecto. En particular, la ontología de recursos digitales tiene usos posibles en multitud de proyectos. Estos proyectos pueden ser similares a este, como un sistema de recursos digitales con carácter medioambiental (nubes, peces, depuradoras, etc.) o incluso proyectos totalmente diferentes, como una clasificación de documentos en un servidor web.

Las instancias también forman parte de un documento separado a la ontología, permitiendo la reutilización de este proyecto tal cual está en otros ámbitos geográficos o utilizando sólo una parte de los recursos digitales disponibles en la web, restringiéndolos, por ejemplo, a los recursos disponibles en un determinado servidor.

En resumen, la ontología creada es adecuada al proyecto, es fácilmente extensible,

y es reusable en otro tipo de proyectos con mayor o menor similitud a este.

6.2. Consideraciones sobre el prototipo de buscador

El buscador es un prototipo básico. Se compone de lo mínimo necesario para ser considerado un buscador: un formulario de búsqueda, una lista de resultados y fichas de puntos de interés.

La aplicación se ha enfocado hacia los puntos de interés en lugar de hacia los recursos digitales, por considerar que los PdI tienen mayor entidad desde el punto de vista de alguien que busque información turística. Orientarlo hacia una simple búsqueda de recursos digitales restaría importancia al término "turístico" que preside el título del proyecto.

Aún siendo un ejemplo muy simple, muestra unas pocas funcionalidades interesantes que sirven como ejemplo de cómo trabajar con una ontología. Por ello es importante observar esta aplicación como un ejemplo tanto de los logros de la ontología creada, como de el cómo crear código para trabajar con ella en Java.

La aplicación separa la carga y tratamiento de la ontología en una clase ConectorDatos. En esta clase se puede ver cómo cargar una ontología desde una URL, cómo activar un modelo de inferencia, como realizar consultas SPARQL y el método utilizado para convertir una instancia de nuestra ontología en un objeto de Java.

Las búsquedas por clase y por localizaciones hacen uso de la inferencia y de la definición transitiva de la propiedad estaSituadoEn. Como se comenta en el apartado anterior, el permitir una búsqueda haciendo uso de la inferencia es uno de los objetivos principales de la ontología y por tanto es necesario mostrar esta fortaleza en el prototipo.

En la aplicación no se hace uso de los tipos de información de los recursos digitales. Aunque la ontología está preparada para ello e incluso hay instancias que hacen uso de la funcionalidad, en la aplicación al final no se ha plasmado esta funcionalidad por falta de tiempo. Una mejora inicial de la aplicación debería hacer uso de esta funcionalidad, mostrando qué tipo de información muestran los recursos digitales de los PdI. En particular es interesante indicar qué tipo de información muestran las páginas web y wikis asociadas.

Con todos sus defectos esta aplicación puede servir como base para un punto de información turística en forma de monolitos táctiles para su uso por parte de los turistas de una zona turística. Está claro que habría que hacer ampliaciones a la aplicación, sobre todo en lo referente a la internacionalización de la aplicación, pero las funcionalidades básicas son adecuadas.

Esta aplicación puede extenderse fácilmente creando una herramienta potente para su uso por parte de oficinas de información turística. Las extensiones podrían ser las comentadas en las conclusiones de la ontología, permitiendo la inclusión de fichas de datos y el enlace con sistemas de información sobre eventos y espectáculos.

Otra de las mejoras a la aplicación sería la ampliación de sus funcionalidades, permitiendo mostrar en una ficha los PdI cercanos al PdI mostrado, otros PdI que comparten clasificación y los PdI que contiene o están contenidos en el PdI. También es un a mejor importante la creación de un mapa interactivo que muestre las

posiciones de varios PdI sobre el mapa, permitiendo filtrar por tipo u otros parámetros.

En resumen, aunque la aplicación tiene sus defectos , sigue siendo un buen ejemplo de la potencia de la ontología creada , y permite crecer hacia una aplicación con funcionalidades como las indicadas, que sería útil en un ámbito empresarial o gubernamental.

7. Glosario de términos

A continuación se explican algunos conceptos encontrados en este documento, ordenados por orden alfabético.

- **Agente de Web Semántica:** Aplicación o programa capaz de buscar en la web contenido semántico en diferentes fuentes y procesar la información obtenida para conseguir un resultado complejo. Un ejemplo de agente podría buscar información sobre especialistas médicos de una ciudad, filtrar según horarios disponibles, localizar los más cercanos a paradas de autobús o metro, y devolver una lista ordenada por precio de los especialistas encontrados.
- **Buscador web:** Aplicación web que permite la búsqueda de información a partir de unos determinados parámetros de búsqueda.
- **CGI:** (Common Gateway Interface) Tecnología por la cual se compone una página HTML como resultado de la ejecución de un programa. Esta tecnología se ha utilizado desde los inicios de Internet para producir páginas dinámicas.
- **Clase:** Indica una categoría característica dentro de una ontología. Se relaciona directamente con un concepto de el dominio. Para más información ver el apartado Clases, propiedades e instancias en la página 10
- **Dominio de una ontología:** Indica el ámbito de interés de una ontología. Se refiere a todos los conceptos que forman parte de un determinado área de conocimiento.
- **HTML:** (Hypertext Markup Language) Lenguaje que permite escribir contenidos para la web, indicando el formato que tienen estos contenidos. El lenguaje es procesado por los navegadores web para mostrar un contenido en Internet.
- **Instancia de una clase:** Una instancia representa un objeto específico del tipo indicado por la clase. Por ejemplo la instancia Castillo de Bellver representa un objeto de la clase Monumento. Para más información ver el apartado Clases, propiedades e instancias en la página 10
- **Java:** Lenguaje de programación orientado a objetos.
- **JSF:** (Java Server Faces) Extensión de Java que permite la creación de documentos HTML a partir de información en una aplicación Java. Utiliza la filosofía CGI.
- **Ontología:** Describe formalmente un dominio de interés, dando una lista finita de términos y su relación entre ellos. Para más información ver el apartado de Ontologías en la página 7
- **OWL:** Lenguaje que extiende el modelo de datos RDF-RDFS para dotarlo de mayor expresividad semántica. Para más información ver el apartado OWL en la página 11

- **Punto de Interés:** Indica un lugar, edificio o estructura que tiene un cierto interés turístico. Siempre se refiere a un lugar físico con unas coordenadas geográficas.
- **RDF:** Modelo de datos que permite definir información de manera semántica. Para más información ver el apartado RDF y RDF-Schema en la página 9
- **RDFS:** (RDF-Schema) Permite definir la terminología que se va a utilizar en un determinado contexto. Para más información ver el apartado RDF y RDF-Schema en la página 9
- **Recurso Digital:** Se refiere a un documento en formato digital que puede ser encontrado en Internet. El documento puede ser una imagen, vídeo, texto, etc. y está identificado por su URL en la red.
- **Tomcat:** Servidor de aplicaciones que permite ejecutar aplicaciones web en Java.
- **UML:** (Unified Modeling Language) Lenguaje en forma de esquemas que permite mostrar la estructura de un producto informático y sus procesos internos.
- **Web Semántica:** Conjunto de tecnologías que permiten expresar información en la web de forma que puedan ser entendidas por un ordenador. Para ello hace uso de ontologías. Para más información ver el apartado sobre Web semántica en la página 6
- **XML:** Lenguaje que permite la estructuración de información en base a etiquetas. Para más información ver el apartado XML y XML Schema en la página 8

8. Bibliografía

8.1. Libros

- Dean Allemang; James Hendler (2007). **Semantic Web for the Working Ontologist**: Morgan Kaufman Publishers.
- Grigoris Antoniou; Frank van Harmelen (2008). **A Semantic Web Primer**: The MIT Press.
- Toby Segaran; Colin Evans; Jamie Taylor (2009). **Programming the Semantic Web**: O'Reilly Media.
- Matthew Horridge (2009). **A Practical Guide To Building OWL Ontologies Using Protégé 4 and CO-ODE Tools**: The University of Manchester.
- Miguel Ángel Abián (2005). **El futuro de la Web**.

8.2. Páginas Web

- Página de Wikipedia sobre Web Semántica: [\[wikipedia\]](#)
- Página del proyecto Jena: <http://jena.sourceforge.net/>
- Tutorial sobre SPARQL: <http://openjena.org/ARQ/Tutorial/index.html>
- Uso de Reasoners en Jena: [\[Reasoners\]](#)
- OWL (Web Ontology Language): [\[OWL\]](#)
- DAML Ontology Library: <http://www.daml.org/ontologies/>
- Protege Ontology Library: [\[Protege\]](#)
- Swoogle – Buscador de ontologías : <http://swoogle.umbc.edu/>

9. Anexos

9.1. Anexo I. Enunciado del proyecto

Títol: "Ontologies turístiques geogràfiques"

L'objectiu d'aquesta proposta és obtenir una ontologia (o model conceptual) que permeti etiquetar fitxers que continguin informació sobre punts d'interès turístic. L'ontologia ha de permetre representar els aspectes relacionats amb l'àmbit turístic (tipus de punt d'interès -monument, església, etc.-, any de creació, estil, ...) i haurà de permetre definir també la latitud i longitud de la informació associada per cadascun dels fitxers.

Per desenvolupar aquest projecte l'estudiant haurà de realitzar les següents tasques:

1) Estudi de l'existència d'ontologies d'informació turística. S'haurà d'estudiar si existeixen ontologies d'aquest tipus i, en cas afirmatiu, si es poden aprofitar en el nostre context.

2) Identificació de la informació a representar i creació de l'ontologia associada.

3) Instància de l'ontologia creada: es crearan instàncies concretes de l'ontologia que descriguin la informació representada en un conjunt de fitxers que es facilitaran a mitjans del TFC. Aquests fitxers contindran informació sobre punts d'interès turístics.

4) Creació d'un prototip simple per avaluar l'ontologia plantejada. L'objectiu del prototip serà identificar quins fitxers es poden servir a l'usuari en funció de les seves preferències.

El funcionament del prototip serà el següent:

1) El prototip haurà d'interactuar amb l'usuari i fer-li preguntes per saber el tipus d'informació turística en que està interessat (museus, esglésies, ...), el format preferit (text, àudio, vídeo, ...) i la seva zona geogràfica.

2) Posteriorment el prototip consultarà les instàncies de l'ontologia creada per trobar els fitxers que satisfacin els criteris de cerca de l'usuari (s'aconsella fer servir una base de dades XML per emmagatzemar les instàncies de l'ontologia),

3) Finalment, el prototip utilitzarà la informació extreta de l'ontologia per crear una pàgina web amb links als fitxers que poden ser d'interès per a l'usuari.

Els fitxers per fer servir es proporcionaran a l'estudiant a mitjans del TFC i estaran disponibles des d'un lloc web allotjat a un servidor de la UOC.