

WEB SEMÁNTICA

SMARTSEVILLA



Bouyich España, Carolina
Calero Pereira, Yanira
Caraza Luis, Jaime
Palomino Sánchez, Sergio
Pérez Melero, Jesús

v130317
v130318
v130263
v130297
v130074

Grupo 10

ÍNDICE

- 1) Obtención de datos.
- 2) Análisis de los datos.
- 3) Modificaciones en los Datos.
- 4) Ontologías Desarrolladas.
- 5) Enlazado de Datos. Valor añadido.
Demostración.
- 6) Conclusión.

1.- Obtención de datos

- Datos: ciudad española de **Sevilla**
- Formato CSV
- Portal de datos abiertos de la ciudad
- Licencia: CC BY 4.0



Datos encontrados:

- Agencia de autobuses de Sevilla
- Calles de Sevilla
- Líneas de autobús y paradas de Sevilla
- Horas de autobuses
- **Museos de Sevilla**
- Bibliotecas de Sevilla
- **Sitios de interés de Sevilla**

2.- Análisis de los datos

- Formato de los datos: especificación general de feeds de *Google Transit* (datasets de transporte)

agency.txt



Archivo: Obligatorio

Nombre del campo	Obligatorio	Detalles
agency_id	Opcional	El campo agency_id es un ID que identifica de forma exclusiva a una empresa de transporte público. Un feed de transporte público puede representar datos de más de una empresa. El agency_id es un conjunto de datos único. Este campo es opcional para los feeds de transporte público que solo incluyan datos de una sola empresa.
agency_name	Obligatorio	El campo agency_name contiene el nombre completo de la empresa de transporte público. Google Maps mostrará este nombre.
agency_url	Obligatorio	El campo agency_url contiene la URL de la empresa de transporte público. El valor debe ser una URL completa que incluya http:// o https:// , y cualquier carácter especial que incluya la URL debe tener el formato de escape correcto. Consulta http://www.w3.org/Addressing/URL/4_URI_Recommentations.html para obtener una descripción de cómo crear valores de URL completas que cumplan los requisitos.

2.- Análisis de los datos (II)

Para cada dataset se realizó un análisis/clasificación, según la siguiente tabla

Column	Type	Comments/Range (rounded)	Problems
--------	------	--------------------------	----------

Resource naming strategy:

- Para términos ontológicos: <http://smartsevilla.es/ontology#>
- Para términos individuales: <http://smartsevilla.es/resources/>

3.- Modificaciones en los Datos

Las modificaciones en los datasets se han realizado mediante las herramientas LODRefine y Notepad++.

- Eliminando columnas en blanco
- Creando nuevas columnas en función de nuestras necesidades.
- Eliminando acentos y caracteres que pudiesen originar problemas.

Exportamos tanto las modificaciones (en JSON) como los RDF generados en TTL.



3.- Modificaciones en los Datos

▼ All	▼ stop_id	▼ stop_code	▼ stop_name	▼ stop_desc	▼ stop_lat	▼ stop_lon	▼ zone_id	▼ stop_url	▼ location_type	▼ parent_station
☆	1.	1	Campana (Sierpes)		37.392658	-5.9953113				
☆	2.	2	Menéndez Pelayo (Puerta Carmona)		37.388371	-5.9849925				
☆	3.	3	Menéndez Pelayo (Puerta de La Carne)		37.386154	-5.9860196				
☆	4.	4	Menéndez Pelayo (Juzgados)		37.382858	-5.9877234				
☆	5.	5	Avda. MarÃa Luisa (Teatro Lope de Vega)		37.378342	-5.9901185				
☆	6.	6	Avda. Parsi (Menorca)		37.37949	-5.9190974				
☆	7.	7	Plaza del Duque (La Campana)		37.39267	-5.9951291				
☆	8.	8	Calatrava (Alameda de Hércules)		37.401508	-5.9940667				
☆	9.	9	Plaza Ponce de LeÃn		37.393009	-5.9876027				
☆	10.	10	PagÃes del Corro (Troya)		37.38118	-6.0005012				

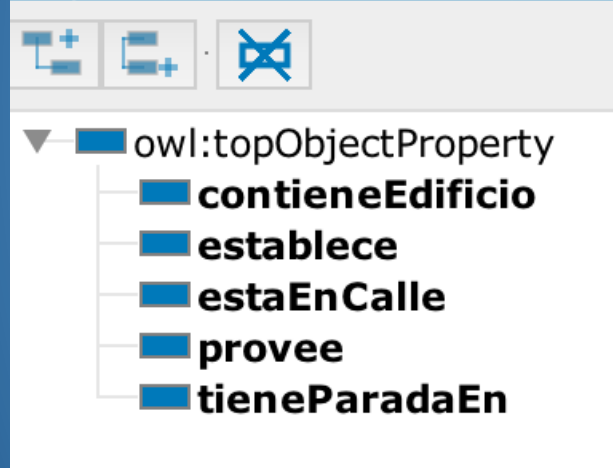


▼ All		▼ stop_id	▼ stop_name	▼ stop_lat	▼ stop_lon	
☆	🔊	1.	1	Campana	37.392658	-5.9953113
☆	🔊	2.	2	Menendez Pelayo	37.388371	-5.9849925
☆	🔊	3.	3	Menendez Pelayo	37.386154	-5.9860196
☆	🔊	4.	4	Menendez Pelayo	37.382858	-5.9877234
☆	🔊	5.	5	Avda. Maria Luisa	37.378342	-5.9901185
☆	🔊	6.	6	Avda. Parsi	37.37949	-5.9190974
☆	🔊	7.	7	Plaza del Duque	37.39267	-5.9951291
☆	🔊	8.	8	Calatrava	37.401508	-5.9940667
☆	🔊	9.	9	Plaza Ponce de Leon	37.393009	-5.9876027
☆	🔊	10.	10	Pages del Corro	37.38118	-6.0005012

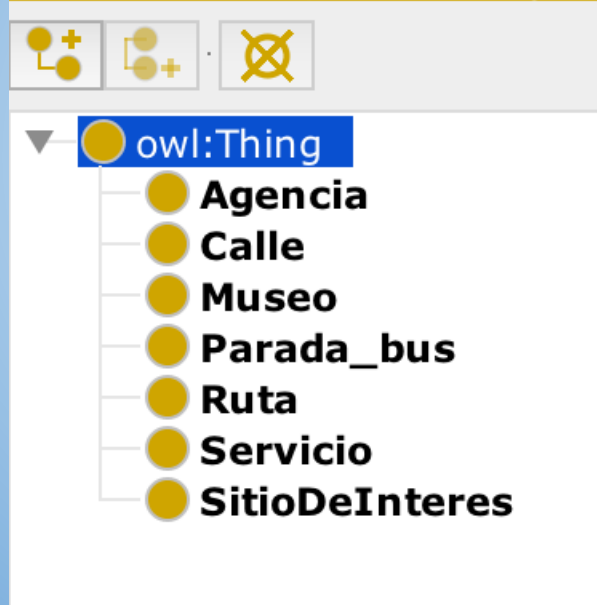
4.- Ontologías Desarrolladas

Clases, Propiedades de Datos y de Objetos de la **primera ontología** desarrollada. Todas las ontologías han sido desarrolladas con Protégé.

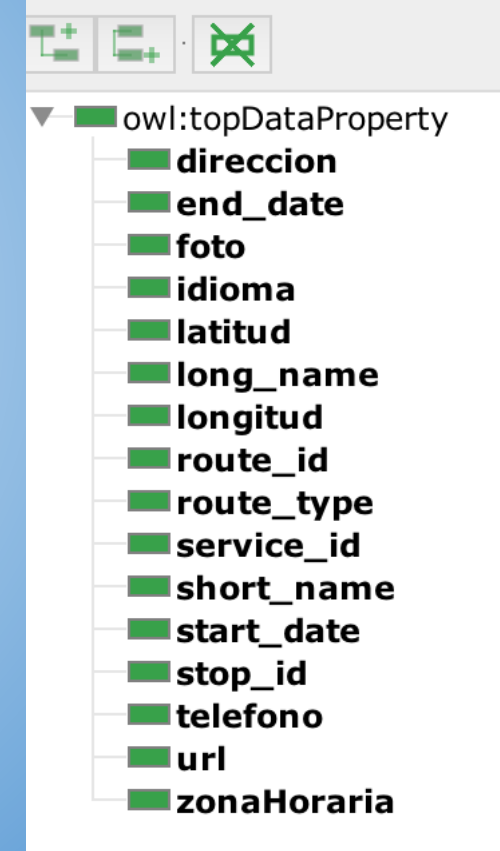
Object property hierarchy:



Class hierarchy: owl:Thing



Data property hierarchy:



4.- Ontologías Desarrolladas (II)

La segunda ontología se desarrolló utilizando información ya existente en **DBpedia**.

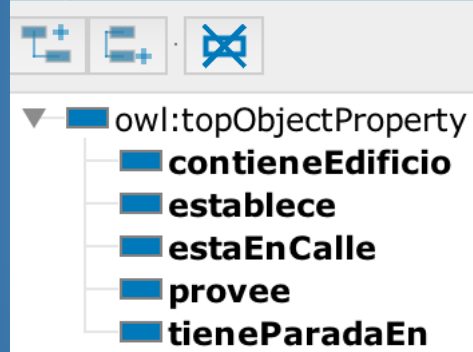
Motivaciones:

- Se adaptaba bien a nuestro modelo.
- Vista y explicada en clase.

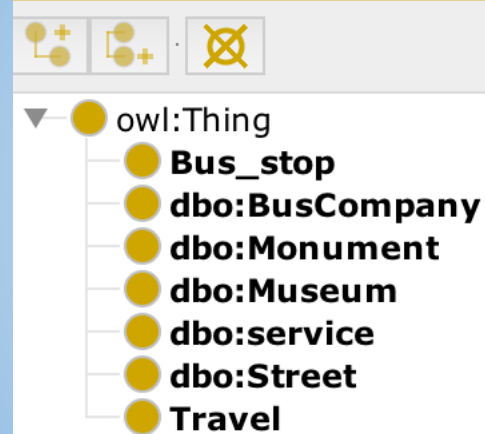


4.- Ontologías Desarrolladas (III)

Object property hierarchy:



Class hierarchy:



Data property hierarchy:



4.- Ontologías Desarrolladas (IV)

Los ficheros RDF se volvieron a generar acorde a la nueva ontología usando de nuevo LODRefine

```
<http://smartsevilla.es/resources/agencias/TUSSAM> a g10:Agencia ;  
  g10:short_name "TUSSAM"^^<xsd:string> ;  
  g10:long_name "Transportes Urbanos de Sevilla S.A.M."^^<xsd:string> ;  
  g10:url <http://www.tussam.es> ;  
  g10:zonaHoraria "Europe/Madrid"^^<xsd:language> ;  
  g10:idioma "es"@es ;  
  g10:telefono "+34 955 479 000"^^<xsd:string> .
```



```
<http://smartsevilla.es/resources/busCompany/TUSSAM> a dbo:BusCompany ;  
  dbp:name "TUSSAM"^^<xsd:string> ;  
  dbp:fullname "Transportes Urbanos de Sevilla S.A.M."^^<xsd:string> ;  
  dbp:url <http://www.tussam.es> ;  
  dbp:language "es"@es ;  
  dbp:phoneNumber "+34 955 479 000"^^<xsd:string> ;
```

5.- Enlazado de datos

Idea Inicial: Ser capaces de enlazar la información de nuestros 8 datasets.

Realidad: Para algunos de los datasets no se obtuvo ningún enlace (relativos a líneas de transporte y calendarios)



Para buscar los enlaces usamos la herramienta **LODRefine** y buscamos coincidencias sobre **esDBpedia** ya que consideramos que era la opción **perfecta**.

5.1- Valor Añadido

Nuestra aplicación:

- **Utiliza los archivos RDF** (TTL) para hallar todo tipo de Información referente a los temas que tratamos
 - Dicha información puede encontrarse en el RDF (especialmente en museos).
 - También **obtiene información enlazando con esDBpedia** para conseguir las descripciones y las imágenes de los museos y monumentos de la ciudad.
- **Por tanto**, la aplicación tiene como fuente principal de información **pequeños ficheros** (13KB y 8 KB) y el hecho de usar datos enlazados nos permite obtener grandes cantidades de información, así como rutas agrupadas por categorías por la ciudad de Sevilla de forma sencilla y eficiente.

5.2- Demostración



