

Práctica 9 - Tripletas RDF

Hugo Pérez Fernández. U0250708@uniovi.es

Sistemas de Información para la Web. Grado de Ingeniería Informática. EII.
universidad de Oviedo. Campus de los Catalanes. Oviedo.

1 Introducción

2 Obtención manual de información estructurada

Primero obtendremos toda la información estructurada de los textos indicados en el Anexo.4.1, usando para ello la web Schema.org:

Texto 1: En el caso de este texto se esta mencionando a una persona (*schema:Person*), cuyo nombre es Miles Davis, ademas esta persona tiene como nacionalidad *schema:Nacionality* Americano (Estadounidense), y su profesión (*schema:Ocupation*) es músico concretamente de Jazz.

Texto 2: En en este caso la información se puede estructurar de la siguiente forma:

- Hay y una persona (*schema:Person*) que se llama Barack Obama cuya ocupación (*schema:Ocupation*) es presidente.
- Hay una organizacion (*schema:Organization*) que se llama Unión Europea.
- Hay una ciudad (*schema:City*) que se llama Washington.
- Hay una tipo de cambio de moneda (*schema:ExchangeRateSpecification*) que se llama Euro y tiene un valor de 1.3.
- Hay una reunion (*schema:BusinessEvent*) entre Barack Obama y la Unión Europea sobre el el valor del tipo de cambio del Euro.

Texto 3: En este último se puede estructura la información del siguiente modo:

- Hay una organización (*schema:Organization*) que se llama The New York Times.
- Hay una persona (*schema:Person*) que se llama Jhon McCarthy.
- Jhon McCarthy está muerto.
- El New York Times informa (*schema:Report*) que Jhon McCarthy esta muerto.
- LISP es un lenguaje de programación.
- John McCarthy inventó LISP.

Con los datos obtenidos anteriormente se, usando la notacion Turtle, la información estructurada queda de la siguiente manera:

Listing 1.1. Archivo Turtle con la información estructurada de los textos.

```
1 @prefix wikidata: <http://www.wikidata.org/> .
2 @prefix foaf: <http://xmlns.com/foaf/0.1/> .
3
4 # Texto 1:
5
6 # Persona de nombre Miles Davis y ocupación músico de
   Jazz
7 <#Miles_Davis> wikidata:Q21503252 foaf:Person ;
8               foaf:name "Miles Davis" ;
9               wikidata:Property:P27 wikidata:Q30 ;
10              wikidata:Property:P106 wikidata:
               Q15981151 .
11
12 # Texto 2:
13
14 # Persona Barack Obama Presidente
15 <#Barack-Obama> wikidata:Q21503252 foaf:Person ;
16               foaf:name "Barack Obama" ;
17               wikidata:Property:P39 wikidata:Q11696
               .
18
19 # Organización llamada Unión Europea
20 <#UE> wikidata:Q21503252 wikidata:Q43229 ;
21      wikidata:Q82799 "Unión Europea" .
22
23 # Ciudad llamada Washington
24 <#WDC> wikidata:Q21503252 wikidata:Q515 ;
25      wikidata:Q82799 "Washington" .
26
27 # Tipo de cambio del euro a 1.3
28 <#Tipo_euro> wikidata:Q21503252 wikidata:Q4916 ;
29             wikidata:Q21503252 wikidata:Q4916 ;
30             wikidata:P1181 "1.3" .
31
32 # Encuentro entre Unión Europea y Barak Obama
33 <#Meeting> wikidata:Q21503252 wikidata:Q2761147 ;
34           wikidata:Property:P710 wikidata:Q458 ;
35           wikidata:Property:P710 wikidata:Q76 ;
36           wikidata:Property:P921 <#tipo_euro> .
37
38 # Texto 3:
39
40 # Organizacion New York Times
41 <#New_York_Times> wikidata:Q21503252 wikidata:Q43229 ;
```

```

42                                wikidata:Q82799 "New York Times" .
43
44 # Persona John McCarthy muerta
45 <#John_McCarthy>      wikidata:Q21503252    foaf:Person ;
46                        foaf:name      "John McCarthy" ;
47                        wikidata:Q21503252    wikidata:Q48422
48
49 # Lenguaje de programación desarrollado por John McCarthy
50 <#LISP> wikidata:Q21503252    wikidata:Q9143 ;
51        wikidata:Q82799 "LISP" ;
52        wikidata:P178 <#John_McCarthy> .
53
54 # The New York Times informa sobre la muerte de John
    McCarthy
55 <#News> wikidata:Q21503252    wikidata:Q732577 ;
56        wikidata:Property:P921 <#John_McCarthy> ;
57        wikidata:Property:P123 <#New_York_Times> .

```

Ahora verificaremos la estructura de la sintaxis RDF de del archivo anterior mediante la herramienta RDFSShape:

RDF Data info

The screenshot shows the RDFSShape tool interface. On the left, under 'Well formed RDF', the input Turtle data is displayed:


```

1 <@prefix wikidata: <http://www.wikidata.org/> .
2 @prefix foaf: <http://xmlns.com/foaf/0.1/> .
3
4 <#UE>      wikidata:Q21503252    wikidata:Q43229 ;
5           wikidata:Q82799      "Unión Europea" .
6
7 <#New_York_Times> wikidata:Q21503252    wikidata:Q43229 ;
8           wikidata:Q82799      "New York Times" .
9
10 <#Tipo_euro> wikidata:P1181 "1.3" ;
11           wikidata:Q21503252    wikidata:Q4916 .
12
13 <#Miles_Davis> wikidata:Property:P106 wikidata:Q15981151 ;
14           wikidata:Property:P27  wikidata:Q30 ;
15           wikidata:Q21503252    foaf:Person ;
16           <rdf:type>             foaf:Person .
    
```

 Below the data, it shows 'Number of statements: 28' and 'DataFormat: turtle'.

On the right, under 'RDF data', the validated data is shown in a similar format, with the same statements. Below this, the 'Data format' is set to 'TURTLE'.

Fig. 1. Resultado de la validación del ttl con RDFSShape.

Una vez verificado podemos traducir el formato turtle a JSONLD, usando para ello la herramienta RDF-Translator, de forma que quedaria de la siguiente manera:

```

1 {
2   "@context": {
3     "foaf": "http://xmlns.com/foaf/0.1/" ,

```

```

4      "rdf": "http://www.w3.org/1999/02/22-rdf-syntax-ns#"
5      "rdfs": "http://www.w3.org/2000/01/rdf-schema#" ,
6      "wikidata": "http://www.wikidata.org/" ,
7      "xsd": "http://www.w3.org/2001/XMLSchema#"
8  },
9  "@graph": [
10     {
11         "@id": "file:///base/data/home/apps/s%7Erdft-
12             translator/2.408516547054015808/#Tipo_euro" ,
13         "wikidata:P1181": "1.3" ,
14         "wikidata:Q21503252": {
15             "@id": "wikidata:Q4916"
16         }
17     },
18     {
19         "@id": "file:///base/data/home/apps/s%7Erdft-
20             translator/2.408516547054015808/#Meeting" ,
21         "wikidata:Property:P710": [
22             {
23                 "@id": "wikidata:Q458"
24             },
25             {
26                 "@id": "wikidata:Q76"
27             }
28         ],
29         "wikidata:Property:P921": {
30             "@id": "file:///base/data/home/apps/s%7Erdft-
31                 translator/2.408516547054015808/#tipo_euro"
32         },
33         "wikidata:Q21503252": {
34             "@id": "wikidata:Q2761147"
35         }
36     },
37     {
38         "@id": "file:///base/data/home/apps/s%7Erdft-
39             translator/2.408516547054015808/#Miles_Davis" ,
40         "foaf:name": "Miles Davis" ,
41         "wikidata:Property:P106": {
42             "@id": "wikidata:Q15981151"
43         },
44         "wikidata:Property:P27": {
45             "@id": "wikidata:Q30"
46         },
47         "wikidata:Q21503252": {

```

```

44         "@id": "foaf:Person"
45     }
46 },
47 {
48     "@id": "file:///base/data/home/apps/s%7Erdf-
        translator/2.408516547054015808/#
        New_York_Times" ,
49     "wikidata:Q21503252": {
50         "@id": "wikidata:Q43229"
51     },
52     "wikidata:Q82799": "New York Times"
53 },
54 {
55     "@id": "file:///base/data/home/apps/s%7Erdf-
        translator/2.408516547054015808/#John_McCarthy"
        ,
56     "foaf:name": "John McCarthy" ,
57     "wikidata:Q21503252": [
58         {
59             "@id": "wikidata:Q48422"
60         },
61         {
62             "@id": "foaf:Person"
63         }
64     ]
65 },
66 {
67     "@id": "file:///base/data/home/apps/s%7Erdf-
        translator/2.408516547054015808/#WDC" ,
68     "wikidata:Q21503252": {
69         "@id": "wikidata:Q515"
70     },
71     "wikidata:Q82799": "Washington"
72 },
73 {
74     "@id": "file:///base/data/home/apps/s%7Erdf-
        translator/2.408516547054015808/#UE" ,
75     "wikidata:Q21503252": {
76         "@id": "wikidata:Q43229"
77     },
78     "wikidata:Q82799": "Unión Europea"
79 },
80 {

```

```

81      "@id": "file:///base/data/home/apps/s%7Erdf-
        translator/2.408516547054015808/#Barack_Obama"
82      ,
83      "foaf:name": "Barack Obama" ,
84      "wikidata:Property:P39": {
85        "@id": "wikidata:Q11696"
86      },
87      "wikidata:Q21503252": {
88        "@id": "foaf:Person"
89      },
90      {
91        "@id": "file:///base/data/home/apps/s%7Erdf-
          translator/2.408516547054015808/#LISP" ,
92        "wikidata:P178": {
93          "@id": "file:///base/data/home/apps/s%7Erdf-
            translator/2.408516547054015808/#
            John_McCarthy"
94        },
95        "wikidata:Q21503252": {
96          "@id": "wikidata:Q9143"
97        },
98        "wikidata:Q82799": "LISP"
99      },
100     {
101       "@id": "file:///base/data/home/apps/s%7Erdf-
        translator/2.408516547054015808/#News" ,
102       "wikidata:Property:P123": {
103         "@id": "file:///base/data/home/apps/s%7Erdf-
          translator/2.408516547054015808/#
          New_York_Times"
104       },
105       "wikidata:Property:P921": {
106         "@id": "file:///base/data/home/apps/s%7Erdf-
          translator/2.408516547054015808/#
          John_McCarthy"
107       },
108       "wikidata:Q21503252": {
109         "@id": "wikidata:Q732577"
110       }
111     }
112   ]
113 }

```

Listing 1.2. Archivo Resultado de la traduccion de turtle a JSONLD.

3 Obtención automática de información estructurada

3.1 Obtención de RDF's.

Obtendremos los documentos RDF para cada texto, par ello usaremos las siguientes herramientas: Open Calais, DBPedia-spotlight y FRED estos se pueden ver correspondientemente en los Anexos: 4.2, 4.2 y 4.2 que posteriormente se han traducido a formato turtle (Anexos: 4.3, 4.3 y 4.3) y a JSONLD (Anexos: 4.4, 4.4 y 4.4).

Dada esta información ahora podremos dar respuesta a las siguientes cuestiones:

¿Qué ontologías usa cada servicio para "tipar" las instancias detectadas en el texto?

¿Existe algún tipo de dichas ontologías que pudiera considerarse equivalente a otro tipo en schema.org? Señala todos los casos que te hayas encontrado con cada servicio, indicando también (si fuera posible) qué propiedades de un tipo mapearían sobre propiedades del otro.

Reflexiona acerca de los motivos que pueden llevar a cada equipo de desarrolladores a producir una ontología propia. Investiga (someramente) sobre el problema de alineación de ontologías.

Utiliza el servicio sameAs.org para localizar equivalencias para todos los tipos fundamentales de Open Calais, DBPedia Spotlight y FRED, así como los homólogos de Schema.org que detectaste. ¿Existe algún servicio para el que no se disponga de información en sameAs.org?

4 Anexos

4.1 Textos de Prueba

Texto 1: "Miles Davis was an american jazz musician."

Texto 2: "President Barack Obama and European Union leaders huddled in Washington amid growing fears over the future of the euro, which closed greater than 1.3 dollars."

Texto 3: "The New York Times reported that John McCarthy died. He invented the programming language LISP."

4.2 Archivos RDF producidos por herramientas

Open Calais

DBPedia-Spotlight

FRED

4.3 Archivos Turtle producidos por herramientas

Open Calais

DBPedia-Spotlight

FRED

4.4 Archivos JSONLD producidos por herramientas

Open Calais

DBPedia-Spotlight

FRED