

Taller de Lenguajes y Tecnologías de la Web Semántica

RDF: Resource Description Framework

Concepción de Sistemas de Información

2021



Contenidos

- Repaso de las Ideas Básicas
- RDF o Cómo Definir un Grafo y Manipularlo.
 - Ideas Básicas
 - Serialización
- Vocabulario

Contenidos

- Repaso de las Ideas Básicas
- RDF o Cómo Definir un Grafo y Manipularlo.
 - Ideas Básicas
 - Serialización
- Vocabulario

Ideas Básicas

- Representar entidades, propiedades y relaciones entre las entidades:
 - La información interesante está en los links.
- Cada recurso se representa por un grafo.
- Cada arco representa alguna forma de relación entre recursos o datos.
- A cada recurso se lo puede clasificar según sus relaciones.

Ideas Básicas

- Se pueden establecer relaciones entre esas clases o entre las relaciones.
- Una película se distribuye países o regiones. Si es una región, se distribuye en cada uno de los países de la región.
- Al tratar todo el grafo, se puede encontrar información que no se conocía. (inferencia)

Un Ejemplo de Recurso



Cómo Trabajar con el Grafo?

- Se necesita un lenguaje que sea adecuado para:
 - Representación de recursos
 - Tratamiento de los mismos
 - Consulta
 - Clasificación
 - Acceso Remoto (Transmisión)

RESOURCE DESCRIPTION FRAMEWORK

Contenidos

- Repaso de las Ideas Básicas
- **RDF o Cómo Definir un Grafo y Manipularlo.**
 - Ideas Básicas
 - Serialización
- Vocabulario

Contenidos

- Repaso de las Ideas Básicas
- RDF o Cómo Definir un Grafo y Manipularlo.
 - Ideas Básicas
 - Serialización
- Vocabulario

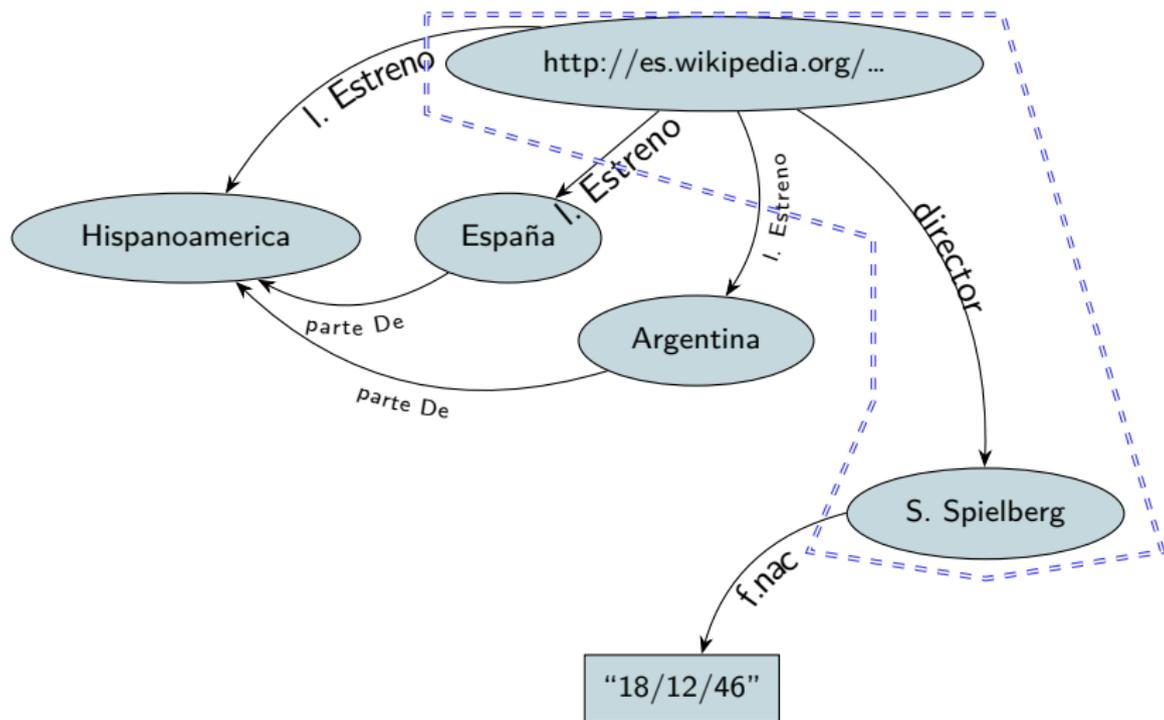
Ideas Básicas

- Todo es un recurso.
 - Tanto los elementos de los que se habla como las cosas que dicen de esos elementos (propiedades) son recursos.
 - Todo recurso está identificado por una URI.
 - No necesariamente una URL ! Podría no ser accesible.

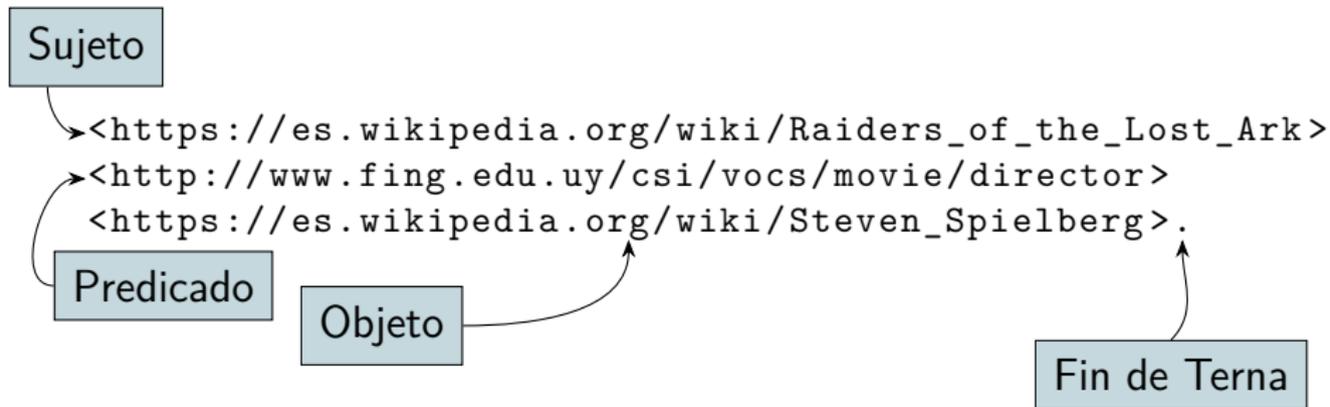
El Modelo de Datos RDF

- Es orientado a grafos.
- Un grafo RDF es:
 - Un Multigrafo.
 - Puede tener muchas aristas, lazos y ciclos sobre sus nodos.
 - Un Grafo Dirigido:
 - Las aristas están orientadas
 - Un Grafo Etiquetado.
 - Cada nodo o arista tiene una etiqueta.
 - La unidad básica es la Terna.
 - Grafo = conjunto de ternas.

Un Grafo



Una Terna



Representa un hecho:

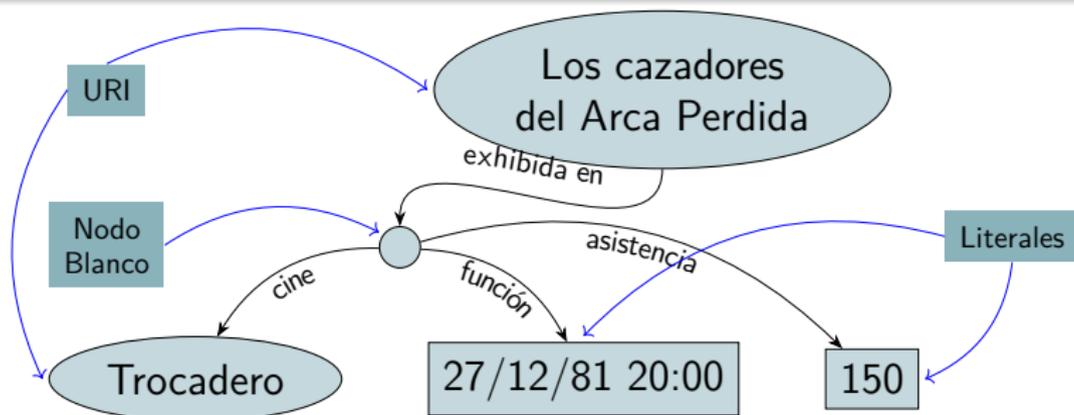
director(Raiders_of_the_Lost_Ark,Steven_Spielberg).

Nodos

Tipos de Nodos

Son tres tipos:

- URIS: Son una referencia a un recurso.
- Literales: son strings con tipos.
- Nodos Blancos: son nodos relativos al grafo.



Las Uris

- Tener muchas URIS, obliga a repetir mucho texto. Solución:
 - Prefijos o Namespaces.
- Los *Namespaces* permiten:
 - Abreviar las URIS.
 - Organizar URIS que están relacionadas.
(Vocabularios)

Ejemplo

```
wes:Los_Cazadores_del_Arca_Perdida
    film:dirigidaPor
        wes:Steven_Spielberg.
wes:Los_Cazadores_del_Arca_Perdida
    film:distribuidaEn
        wes:España.
wes:Los_Cazadores_del_Arca_Perdida
    film:distribuidaEn
        wes:Argentina.
wes:Los_Cazadores_del_Arca_Perdida
    film:distribuidaEn
        wes:Hispanoamerica.
wes:Los_Cazadores_del_Arca_Perdida film:Estreno 1981.
```

- Es normal usar varios prefijos en el mismo grafo.

RDF: Definiciones

Terna RDF

Es una terna $\langle S, P, O \rangle$ en donde:

- S es una URI o un nodo blanco.
- P es una URI
- O es una URI, un nodo blanco o un literal.

Grafo RDF

Conjunto de ternas

Dataset

Es un conjunto de grafos en donde hay un **grafo por defecto** y un **conjunto de Named Graphs**.

Named Graph

Es un grafo RDF con una URI asociada.

- Se considera que está formado por *QUADS* - cuaternas formadas por la URI del grafo y una terna.

Ejemplo

- Ver: http://dbpedia.org/page/Raiders_of_the_Lost_Ark
- Observar:
 - Namespaces (Vocabularios).
 - Recursos.
 - Literales.

Contenidos

- Repaso de las Ideas Básicas
- **RDF o Cómo Definir un Grafo y Manipularlo.**
 - Ideas Básicas
 - **Serialización**
- Vocabulario

Serialización en XML

XML (a)

```
<rdf:RDF>
  <rdf:Description rdf:about='wes:Los_Cazadores_del_Arca_Perdida'>
    <film:director rdf:resource='wes:Steven_Spielberg' />
    <film:distribuidaEn rdf:resource='wes:España' />
    ...
  </rdf:Description>
</rdf:RDF>
```

XML (b)

```
<rdf:RDF>
<rdf:Description rdf:about='wes:Los_Cazadores_del_Arca_Perdida'
  film:director='wes:Steven_Spielberg'>
  <film:distribuidaEn rdf:resource='wes:España' />
</rdf:Description>
</rdf:RDF>
```

Serialización en XML

- Ventajas
 - Existe gran cantidad de software para su tratamiento
- Desventajas
 - Es verboso
 - Admite diferentes serializaciones de formas que pueden complicar el procesamiento.

Serialización en N-Triples

N-Triples

```
<http://es.wikipedia.org/.../Los_Cazadores_del_Arca_Perdida>  
  <http://.../films#dirigidaPor>  
    <http://.../Steven_Spielberg>.  
<http://.../Los_Cazadores_del_Arca_Perdida>  
  <http://.../films#distribuidaEn> <http://.../España> .
```

- Es una terna que termina en "." en donde cada parte está separada por blancos o fin de línea.

Serialización en N-Triples

- Ventajas:
 - Es simple de procesar con casi cualquier lenguaje o paquete.
- Desventajas:
 - No permite abreviar URI's y obliga a repetir todas las partes de la terna.
- Solución:

TURTLE

Serialización en Turtle

TURTLE

```
@prefix wes: <http://es.wikipedia.org/wiki/ > .
@prefix film: <http://www.fing.edu.uy/inco/.../film# > .

wes:Los_Cazadores_del_Arca_Perdida
  film:dirigidaPor wes:Steven_Spielberg ;
  film:distribuidaEn wes:España ,
                    wes:Argentina ,
                    wes:Hispanoamerica ;
  film:Estreno 1981 .
```

- Admite Prefijos.
- Permite compartir sujetos y/o propiedades.

Serialización en Turtle

```
dbpedia-owl:budget rdf:type rdf:Property .
```

```
dbpedia:Raiders_of_the_Lost_Ark
  rdf:type yago:AmericanAdventureFilms,
          dbpedia-owl:Film;
  rdfs:label "'Poszukiwacze zaginionej Arki'"@pl ,
          "'Raiders of the Lost Ark'"@en ;
  dbpedia-owl:budget "'1.8E7'"@ns23:usDollar.
```

Idioma del
String



'@pl

Tipo del
String



'@ns23:usDollar.

Serialización: Resumen

- Hay varias formas:
 - XML
 - N-Triple
 - Turtle
 - Otras (N3, Json, RDFa,...)
- XML y Turtle son recomendación de W3C para serializar RDF.

Contenidos

- Repaso de las Ideas Básicas
- RDF o Cómo Definir un Grafo y Manipularlo.
 - Ideas Básicas
 - Serialización
- Vocabulario

Vocabulario

- Intuitivamente, un **vocabulario** es un conjunto de identificadores de clases y propiedades con una semántica determinada.
- Esa semántica puede estar definida formalmente, o mediante el acuerdo de un grupo de personas que deciden usarla para decir lo mismo.

Vocabulario RDF

Propiedad que dice
en qué clase está
el sujeto

Clase de las
Propiedades

`dbpedia-owl:budget` `rdf:type` `rdf:Property`.

```
dbpedia:Raiders_of_the_Lost_Ark
  rdf:type yago:AmericanAdventureFilms,
          dbpedia-owl:Film;
  rdfs:label "'Poszukiwacze zaginionej Arki'"@pl ,
          "'Raiders of the Lost Ark'"@en ;
  dbpedia-owl:budget "'1.8E7'"^ns23:usDollar.
```

Otros Vocabularios

- Foaf:
Friend of a Friend se utiliza para describir personas y/o organizaciones.
<http://xmlns.com/foaf/0.1/>
- Rdfs:
 - RDF-Schema se utiliza para describir esquemas en rdf. (se va ver en más detalle)
<http://www.w3.org/2000/01/rdf-schema#>
- DC:
 - Dublin Core es un vocabulario para especificar recursos.
<http://purl.org/dc/elements/1.1/>

Bibliografía

- Cyganiak, R., , Wood, D. (2012). RDF 1.1 Concepts and Abstract Syntax (W3C Working Draft). W3C. Retrieved from <http://www.w3.org/TR/rdf11-concepts/>
- Gandon, F. L., Krummenacher, R., Han, S.-K., , Toma, I. (2011). Semantic Annotation and Retrieval: RDF. In J. Domingue, D. Fensel, , J. A. Hendler (Eds.), *Handbook of Semantic Web Technologies* (pp. 117–155). Springer Berlin Heidelberg. Retrieved from http://dx.doi.org/10.1007/978-3-540-92913-0_4
- Hitzler, P., Krötzsch, M., , Rudolph, S. (2009). *Foundations of Semantic Web Technologies*. Chapman , Hall/CRC.