

# Introducción a RDF

---

Lenguajes y Tecnologías de la Web Semántica

## Ideas Básicas de la Web Semántica

---

- 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.
- 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)

# Recurso: Los Cazadores del Arca Perdida

## Lugares de Distribución

## Tipo

### Raiders of the Lost Ark

*Raiders of the lost ark* (conocida también como *Indiana Jones and the raiders of the lost ark*, en [España](#) y [Argentina](#) como *En busca del arca perdida* y en el resto de [Hispanoamérica](#) como *Los cazadores del arca perdida*) es una [película de aventuras](#) estadounidense dirigida por [Steven Spielberg](#), estrenada en [1981](#) y producida por [George Lucas](#). Es el primer filme en ser estrenado de la franquicia de *Indiana Jones* y el segundo en [términos cronológicos de la serie](#). Obtuvo cinco [premios Óscar](#), de un total de nueve nominaciones (que inclúan la categoría de "mejor película").

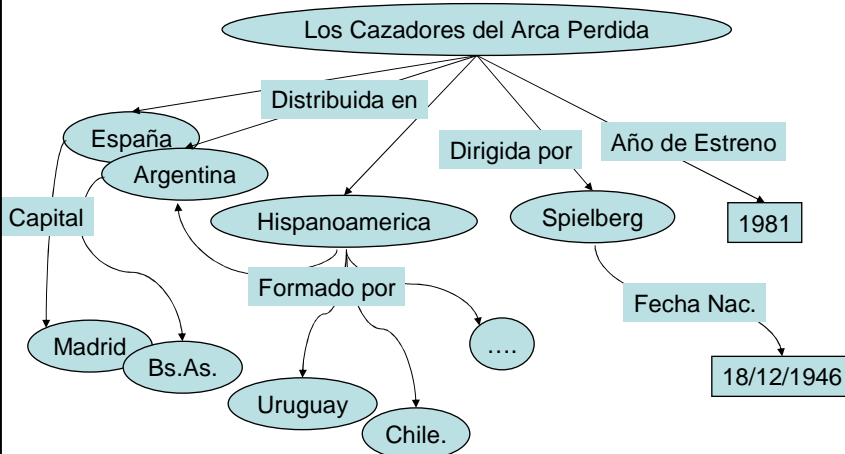
El [Origen](#) ito por [Lawrence](#) [Dirigida Por](#) a cabo [Año de Estreno](#) e cuatro días en la cual Spielberg, Lucas y el [denimieron](#) los principales [e](#) [ivos](#) de la trama. El reparto estuvo integrado por [Harrison Ford](#), [Karen Allen](#), [Paul Freeman](#), [John Rhys-Davies](#) y [Denholm Elliott](#); el [rodaje](#) fue realizado en [Inglaterra](#). La [Rechella](#)

### Raiders of the lost ark



La búsqueda del Arca de la Alianza es el tema central de

# Representación



## 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

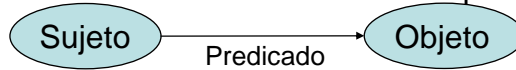
## Ideas Básicas detrás de RDF

---

- 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.
- La unidad básica es la Terna
  - <sujeito> <predicado> <objeto>
  - Da el valor de una propiedad a un sujeto.
- Cualquier grafo se puede ver como un conjunto de ternas

# El Modelo de Datos de 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.

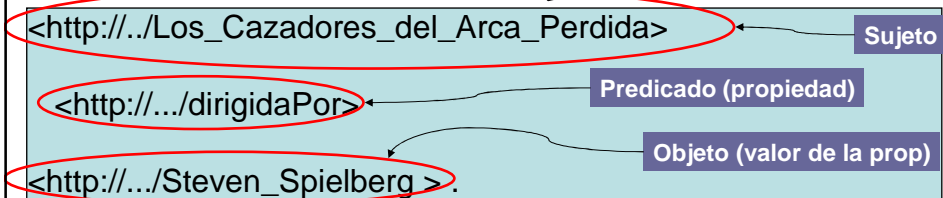
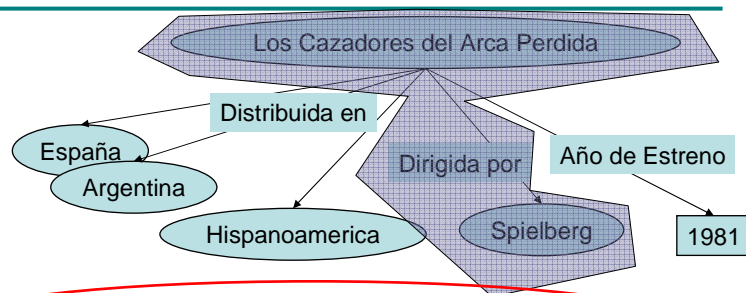


C.S.I.

Experimentos con la Web Semántica: RDF

7

## Ejemplo 1



C.S.I.

Experimentos con la Web Semántica: RDF

8

## Problema: Las Uris

- Deberían estar organizadas de alguna forma “comprensible”.
  - Si es para películas: .../film/Los\_Cazador...
  - Si es para personas:.../personas/Steven\_Sp...
  - Esto no siempre se cumple.
- Se pueden organizar usando la idea de Namespaces.

## Ejemplo: RDF

Nombre	Definición
wes:	<a href="http://es.wikipedia.org/wiki/">http://es.wikipedia.org/wiki/</a>
film:	<a href="http://www.fing.edu.uy/inco/grupos/csi/ontologies/films#">http://www.fing.edu.uy/inco/grupos/csi/ontologies/films#</a>

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:AñoEstreno 1981.

## Namespaces y Vocabularios

---

- Semántica = Significado conocido.
- Los términos con significados acordados se agrupan en vocabularios identificados por un namespace.

Nombre	Definición
rdf	<a href="http://www.w3.org/1999/02/22-rdf-syntax-ns#">http://www.w3.org/1999/02/22-rdf-syntax-ns#</a>
rdfs	<a href="http://www.w3.org/2000/01/rdf-schema#">http://www.w3.org/2000/01/rdf-schema#</a>
movie	<a href="http://data.linkedmdb.org/resource/movie">http://data.linkedmdb.org/resource/movie</a>
foaf	<a href="http://xmlns.com/foaf/0.1/">http://xmlns.com/foaf/0.1/</a>

## Vocabularios

---

- Un Vocabulario es un conjunto de términos con un significado conocido que se utilizan para describir recursos.
- Tiene un namespace asociado, público y conocido.
- **CONCLUSION:**
  - Si se define un vocabulario, debe ser conocido por quienes necesiten utilizarlo.

## Serialización o Cómo Escribir RDF

- XML

```
...
  <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>

```

---

```
...
  <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>

```

## XML

- Ventajas

- Ampliamente usado por lo que existe mucho software para su tratamiento.

- Desventajas

- Verboso
- Representa árboles y no grafos, por lo que requiere codificación de estructuras.
- Admite diferentes serializaciones

## Serialización en N-Triple

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

- Una lista de ternas terminadas en “.”
- Ventajas:
  - Simple y claro.
- Desventajas:
  - Repite mucho.
  - No admite namespaces.

## Serialización en Turtle

```
Wes:Los_Cazadores_del_Arca_Perdida film:dirigidaPor  
wes:Steven_Spielberg;  
  film:disribuidaEn wes:España, wes:Argentina,  
wes:Hispanoamerica;  
  film:AñoEstreno 1981.
```

- Misma idea que NTriple, agregando sintaxis para no repetir.
  - “;” comparte sujeto
  - “,” comparte sujeto y propiedad
  - Termina en .



## Serialización en Turtle

---

- Admite el uso de namespaces.

```
@prefix wes: < http://es.wikipedia.org/wiki/ > .
@prefix film:
<http://www.fing.edu.uy/inco/grupos/csi/ontologies/film# > .
wes:Los_Cazadores_del_Arca_Perdida
    film:dirigidaPor wes:Steven_Spielberg ;
    film:distribuidaEn wes:España ,
                        wes:Argentina ,
                        wes:Hispanoamerica ;
    film:AñoEstreno 1981 .
```

## Serialización: Resumen

---

- Hay varias formas:
  - XML
  - N-Triple
  - Turtle
  - Otras (N3, Json, RDFa,...)
- XML es la única que hasta el momento es recomendación W3C.
- Turtle está como Candidata (19/2/2013) pero es ampliamente usado.

## RDF: Estado

---

- Recomendación de W3C desde 2004.
- Versión 1.1 en draft.
  - Se trata de mejorar la definición de ciertos elementos como por ejemplo, qué es un documento RDF, qué es un Dataset, etc.
- Namespace:
  - Rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>

## 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.

## Nodos Blancos



## Ejemplo

- Ver:
  - [http://dbpedia.org/page/Raiders of the Lost Ark](http://dbpedia.org/page/Raiders_of_the_Lost_Ark)
- Observar:
  - Namespaces (Vocabularios)
  - Recursos y literales.

## Elementos del Vocabulario RDF

Propiedad que dice en qué clase está el sujeto

bbpedia-owl:budget

Clase de las propiedades

**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

C.S.I.

Experimentos con la Web Semántica: RDF

23

Tipo

## Explotando RDF: Introducción a Sparql

PREFIX rdfs: <http://www.w3.org/2000/01/

PREFIX dbpprop: <http://dbpedia.org/prop

PREFIX dbpedia-owl: <http://dbpedia.org/c

PREFIX dbpedia: <http://dbpedia.org/resou

El resultado surge de todos los subgrafos que cumplen con el patrón.

select ?s ?p ?o

where {

?s rdfs:label "Raiders of the Lost Ark"@en

?s ?p ?o

}

Variables

El resultado son todas las ternas que cumplen con el patrón.

Patrón de Grafo

C.S.I.

Experimentos con la Web Semántica: RDF

24

## Introducción a Sparql.

---

- La expresión “select” devuelve una tabla con los valores de las variables.
- Se aplica sobre un grafo por defecto, o se puede especificar “from”.

## 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

---

- Carroll, J. J., & Klyne, G. (2004). *Resource Description Framework (RDF): Concepts and Abstract Syntax* (W3C Recommendation). W3C. Retrieved from <http://www.w3.org/TR/rdf-concepts/>
- 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](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.