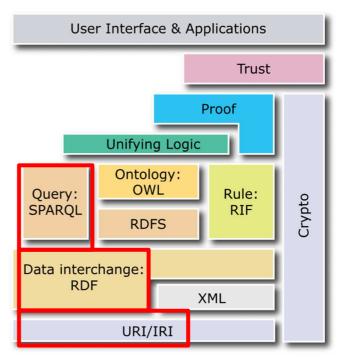
Resource Description Framework (RDF)

José Samos (<u>isamos@ugr.es</u>) Departamento de LSI, UGR

Granada, 11 de Enero de 2019

Contenido

- Datos, Metadatos e Identificadores en la Web
- El modelo de RDF
- La sintaxis de RDF en XML
- Consultas en SPARQL
- Conclusiones y Bibliografía



Datos, Metadatos e Identificadores en la Web

Anidamiento de marcas en XML

¡Anidamiento opuesto para expresar la misma información! No tiene un significado estándar, depende de la aplicación.

Lo que XML nos ofrece y de lo que carece

```
<music genre="classical">
    <title>Eine Kleine Nacht Muzik</title>
    <composer>Mozart</composer>
    <key>E Flat</key>
    <tempo>2/4</tempo>
</music>

<document type="classical music">
         <name>Eine Kleine Nacht Muzik</name>
         <author>Mozart</author>
         </document>
```

Nos ofrece estructura, facilidad de manipulación de datos.

(Avance respecto a HTML)

"composer" = "author" = "creator" = ...

Carece de **significado**.



Usamos un **identificador único** para nombrar cada propiedad:

http://purl.org/dc/elements/1.1/Creator, ...

5

Dublin Core (Dublin, Ohio, 1995)



```
<HTMT.>
  <HEAD>
     <TITLE>I will stand says Portillo</TITLE>
     <META NAME="DC.Title" CONTENT="I will stand says Portillo">
     <META NAME="DC.Creator" CONTENT="Craig Hoy">
     <META NAME="DC.Subject" CONTENT="Tory leadership contest">
  </HEAD>
  <BODY>
     Michael Portillo has announced that he will stand for the leadership
     of the Conservative Party.
     Launching his bid for the Tory crown, Portillo said the Conservative
     Party had to seriously review its approach.
     "Unless the Conservative Party makes huge changes in its style and in
     the issues on which it focuses the party could slip still further in
     public respect," he said.
  </BODY>
</HTML>
```

Se determinó un conjunto de propiedades comunes a la mayoría de tipos de documentos.

Al conjunto de propiedades se le llama *Dublin Core* (DC) initiative.

Conjunto de 15 elementos de información para describir fuentes de información: Contributor, Coverage, Creator, Date, Description, Format, Identifier, Language, Publisher, Relation, Rights, Source, Subjet, Title y Type.

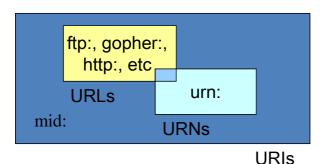
Identificadores y taxonomías

Usamos un **identificador único** para nombrar cada valor del conjunto de valores posible (taxonomía): http://taxonomies.org/People/Mozart, ...

- Taxonomía (o vocabulario controlado): Conjunto de palabras (a menudo sobre un tema) entre las que escogemos para especificar propiedades.
- Identificamos a los elementos que aparecen en páginas distintas de forma única, los interrelacionamos, **tratamos la Web como si fuera una BD**.

Uniform Resource Identifier (URI)

- Especifican el objeto sobre el que hablamos.
- URI: Una forma uniforme de identificar recursos. Podemos asignar un URI a cualquier cosa; cualquier cosa que tenga un URI asignado decimos que "está en la Web". "The Web is an information space. URIs are the points in that space."
 Sólo identifica, no localiza recursos.
- Internationalized Resource Identifier (IRI), extienden la sintaxis de URIs para usar más caracteres.
- Los URIs están descentralizados, ninguna organización los controla (algunos sí, como los del esquema http: dependen de sistemas centralizados –DNS).



Problema: **Múltiples URIs pueden** representar al mismo elemento.

http://www.w3.org/Addressing/

Datos y Metadatos

http://www.ePolitix.com/Articles/0000005a4787.htm

```
<HTML>
  <HEAD>
    <TITLE>I will stand says Portillo</TITLE>
    <META NAME="DC.Title" CONTENT="I will stand says Portillo">
    <META NAME="DC.Creator" CONTENT="Craig Hoy">
    <META NAME="DC.Subject" CONTENT="Tory leadership contest">
  </HEAD>
                                                                     XML
          <document
  <BODY>
            type="News Item"
  </BODY>
           url="http://www.ePolitix.com/Articles/0000005a4787.htm"
</HTML>
           xmlns:dc="http://purl.org/dc/elements/1.1/">
              <dc:Title>I will stand says Portillo</dc:Title>
              <dc:Creator>Craig Hoy</dc:Creator>
              <dc:Subject>Tory leadership contest</dc:Subject>
          </document>
```

Recurso	Propiedad	Valor
http://www.ePolitix.com/	http://purl.org/dc/	I will stand
Articles/000005a4787.htm	elements/1.1/Title	says Portillo
http://www.ePolitix.com/	http://purl.org/dc/	Craig Hoy
Articles/0000005a4787.htm	elements/1.1/Creator	
http://www.ePolitix.com/	http://purl.org/dc/	Tory
Articles/0000005a4787.htm	elements/1.1/Subject	leadership
		contest

9

Metadatos en la Web

B	D	** - 7
Recurso	Propiedad	Valor
http://www.ePolitix.com/	http://purl.org/dc/	I will stand
Articles/0000005a4787.htm	elements/1.1/Title	says Portillo
http://www.ePolitix.com/	http://purl.org/dc/	Craig Hoy
Articles/0000005a4787.htm	elements/1.1/Creator	
http://www.ePolitix.com/	http://purl.org/dc/	Tory
Articles/0000005a4787.htm	elements/1.1/Subject	leadership
Recurso	Propiedad	Valor

Recurso	Propiedad	Valor
http://www.ePolitix.com/	http://purl.org/dc/	I will stand says
Articles/0000005a4787.htm	elements/1.1/Title	Portillo
http://www.ePolitix.com/	http://purl.org/dc/	http://taxonomies.org/
Articles/0000005a4787.htm	elements/1.1/Creator	People/CraigHoy
http://www.ePolitix.com/	http://purl.org/dc/	Tory leadership
Articles/0000005a4787.htm	elements/1.1/Subject	contest

- Los metadatos se pueden expresar como un conjunto de tríos.
- La clave para compartir metadatos en la Web es usar URIs para obtener identificadores únicos.
- El medio para intercambiar, obtener e interpretar fácilmente los datos en la Web: XML.
- ¿Qué estructura han de tener los documentos XML?

El modelo de RDF

Terminología

- La especificación del modelo es independiente de su sintaxis.
- Cada trío es una sentencia (statement): Hacemos sentencias sobre objetos o recursos (resources), se le asocian propiedades con valores, a cada par propiedad-valor se le llama propiedad (property).

Recurso	Propiedad	Valor
http://www.ePolitix.com/	http://purl.org/dc/	Craig Hoy
Articles/0000005a4787.htm	elements/1.1/Creator	

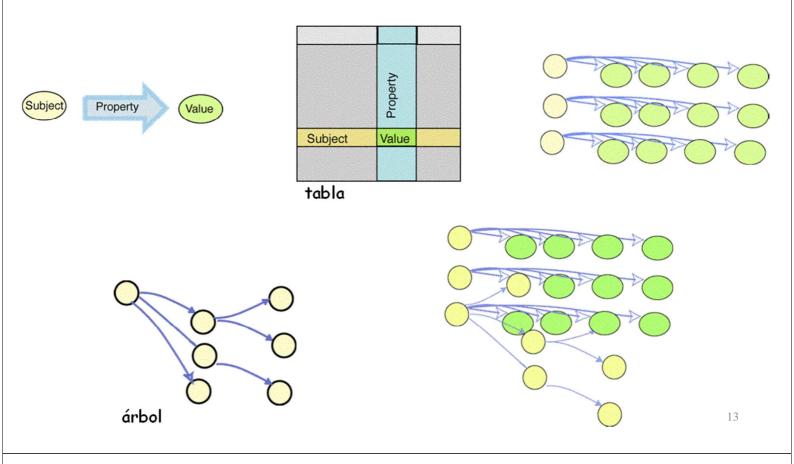
Recurs	80	Propiedad	Valor
http:/	//www.ePolitix.com/	http://purl.org/dc/	http://taxonomies.org/
Artic	es/0000005a4787.htm	elements/1.1/Creator	People/CraigHoy

http://purl.org/dc/elements/1.1/Creator CraigHoy

http://www.ePolitix.com/http://purl.org/dc/elements/1.1/Creator CraigHoy

http://www.ePolitix.com/http://purl.org/dc/elements/1.1/Creator http://taxonomies.org/People/CraigHoy

Representación de datos mediante tríos



Tres vistas de una Sentencia

- Una tripleta: La tripleta (x,P,y) se puede considerar como una fórmula lógica P(x,y). RDF sólo permite predicados binarios.
- Un trozo de un grafo
- Un trozo de código XML



Así, un documento RDF se puede ver como:

- Un conjunto de tripletas (de fórmulas lógicas)
- Un grafo (red semántica)
- Un documento XML

El modelo de RDF es independiente de las vistas.

Formatos de serialización de RDF

- N-Triples (subconjunto de N3)
- N3 (Notation 3 RDF): Tim Berners-Lee (gnames)

N3

- RDF/XML
- RDFa

```
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>.
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>.
@prefix foaf: <http://xmlns.com/foaf/0.1/>.
@prefix doap: <http://usefulinc.com/ns/doap#>.
@prefix dc: <http://purl.org/dc/elements/1.1/>.

@prefix another: <http://swx.org/>.
@prefix i <http://anothercompany.org/>.
@prefix : <http://www.csd.abdn.ac.uk/~ggrimnes/dev/imdb/IMDB#>.
@prefix titles: <http://www.imdb.com/title/>.
```

title:tt0083658 a :Movie ; dc:title "Bladerunner" ;

15

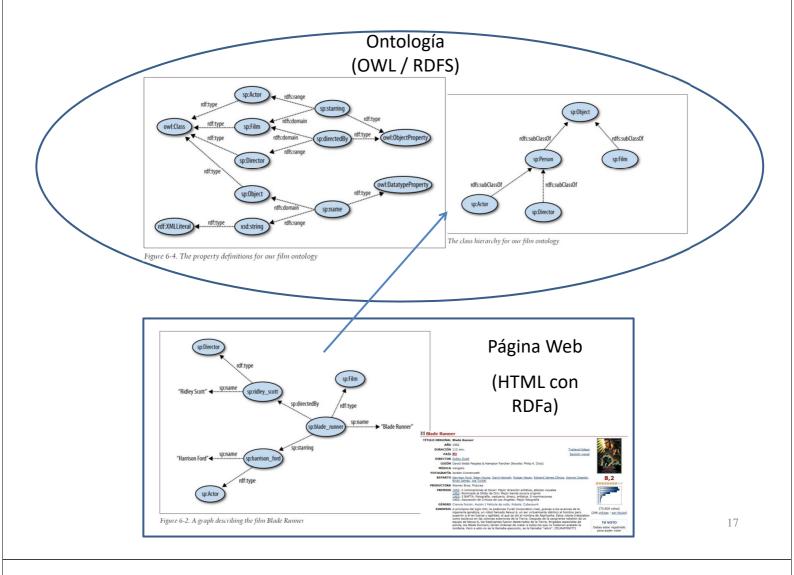
XML

<u>sc</u>°

RDFa

RDFa Primer

Bridging the Human and Data Webs



Sintaxis de RDF en XML

XML y RDF

- XML no forma parte del modelo de datos de RDF.
- Es importante por ser XML el estándar de facto para representar e intercambiar información estructurada en la Web.
- La Web Semántica requiere que las máquinas puedan acceder y procesar la representación que se use, XML cumple esos requisitos.

NO es parte del modelo de datos de RDF.

19

Sintaxis RDF: RDF/XML

Para varias sentencias:

Sentencias en XML

- Un documento RDF se representa mediante un elemento XML con la marca rdf:RDF
- El contenido de este elemento es un conjunto de descripciones que usan la marca rdf:Description
- Cada descripción define una sentencia sobre un recurso, cada recurso se puede identificar de tres formas:
 - Atributo about, referencia un recurso existente.
 - Atributo **ID**, crea un nuevo recurso.
 - Sin un nombre, crea un recurso anónimo.

Varias sentencias con el mismo sujeto

```
xnlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:dc="http://purl.org/metadata/dublin_core#">
 <rdf:Description
 rdf:about="http://www.ePolitix.com/Articles/0000005a4787.htm">
    <dc:Creator>Craig Hoy</dc:Creator>
 </rdf:Description>
 <rdf:Description
 rdf:about="http://www.ePolitix.com/Articles/0000005a4787.htm">
    <dc:Publisher>ePolitix</dc:Publisher>
 </rdf:Description>
                     xnlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
/rdf:RDF>
                     xmlns:dc="http://purl.org/metadata/dublin_core#">
                      <rdf:Description
                       rdf:about="http://www.ePolitix.com/Articles/0000005a4787.htm">
                          <dc:Creator>Craig Hoy</dc:Creator>
                          <dc:Publisher>ePolitix</dc:Publisher>
                       </rdf:Description>
                                  <rdf:RDF
                     </rdf:RDF>
                                   xnlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
                                   xmlns:dc="http://purl.org/metadata/dublin_core#">
                                    <rdf:Description
                                     rdf:about="http://www.ePolitix.com/Articles/0000005a4787.htm"
                                     dc:Creator="Craig Hoy"
                                     dc:Publisher="ePolitix"/>
                                  </rdf:RDF>
```

Sentencias anidadas

```
<rdf:RDF
xnlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:dc="http://purl.org/metadata/dublin_core#"
xmlns:p="http://taxonomies.org/People/Attributes#">
 <rdf:Description
  rdf:about="http://www.ePolitix.com/Articles/0000005a4787.htm">
   <dc:Creator rdf:resource="http://taxonomies.org/People/CraigHoy"/>
 </rdf:Description>
 <rdf:Description
  rdf:about="http://taxonomies.org/People/CraigHoy">
     <p:Name>Craig Hoy</p:Name>
                                 <rdf:RDF
  </rdf:Description>
                                  xnlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
                                  xmlns:dc="http://purl.org/metadata/dublin_core#"
</rdf:RDF>
                                  xmlns:p="http://taxonomies.org/People/Attributes#">
                                   <rdf:Description
                                    rdf:about="http://www.ePolitix.com/Articles/000005a4787.htm">
                                    <dc:Creator>
                                      <rdf:Description
                                       rdf:about="http://taxonomies.org/People/CraigHoy">
                                         <p:Name>Craig Hoy</p:Name>
                                      </rdf:Description>
                                    </dc:Creator>
                                   </rdf:Description>
                                 </rdf:RDF>
```

Atributos *rdf:about* y *rdf:ID*

```
<rdf:RDF>
 <rdf:Description</pre>
  rdf:about="http://www.ePolitix.com/Authors/CraigHoy"/>
    <v:Name>Craig Hoy</v:Name>
    <v:Email>Craig.Hoy@ePolitix.com</v:Email>
 </rdf:Description>
</rdf:RDF>
<rdf:RDF>
                                               http://www.ePolitix.com/RDF/someRDF.rdf
  <rdf:Description</pre>
   rdf:about="/Authors/CraigHoy"/>
                                                 URI relativo, equivalente al anterior si este fragmento
     <v:Name>Craig Hoy</v:Name>
                                                 está almacenado en un documento como el indicado.
     <v:Email>Craig.Hoy@ePolitix.com</v:Email>
  </rdf:Description>
</rdf:RDF>
<rdf:RDF>
                                               http://www.ePolitix.com/RDF/someRDF.rdf
  <rdf:Description</pre>
   rdf:about="/Authors#CraigHoy"/>
                                                 URI relativo y con identificador fragmento.
     <v:Name>Craig Hoy</v:Name>
     <v:Email>Craig.Hoy@ePolitix.com</v:Email>
  </rdf:Description>
</rdf:RDF>
                                               http://www.ePolitix.com/Authors.rdf
<rdf:RDF>
                                                 El atributo Rdf:ID permite abreviar URIs si el recurso
  <rdf:Description</pre>
```

rdf:ID="CraigHoy"/>

</rdf:Description>

</rdf:RDF>

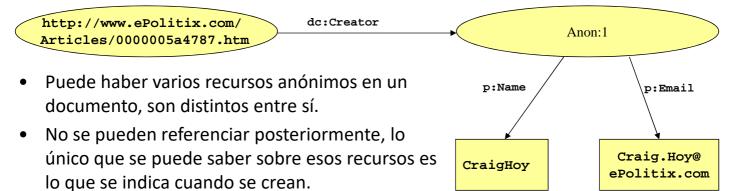
<v:Name>Craig Hoy</v:Name>

<v:Email>Craig.Hoy@ePolitix.com</v:Email>

Existe en ese lugar (ID lo define), los otros no tienen que existir.

referido es el documento que contiene los metadatos.

Recursos anónimos



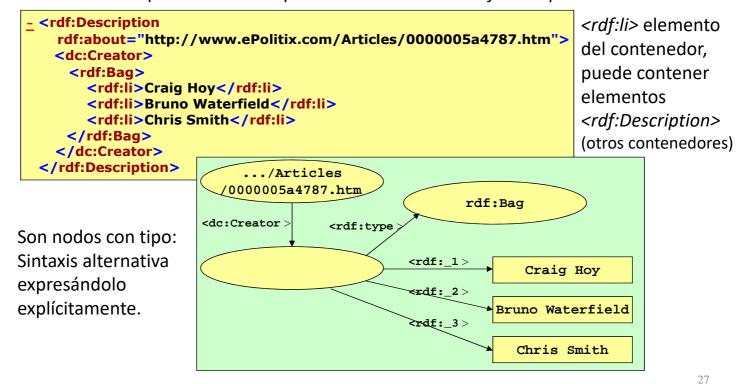
25

Atributo rdf:type y elementos con tipo

```
<rdf:RDF>
 <rdf:Description</pre>
  rdf:about="http://www.ePolitix.com/Authors/CraigHoy"
  rdf:type="http://www.schemas.org/Schemas#Person">
   <dc:Format>text/html</dc:Format>
   <p:Name>Craig Hoy</p:Name>
   <p:Email>Craig.Hoy@ePolitix.com</p:Email>
 </rdf:Description>
                         <rdf:RDF>
</rdf:RDF>
                           <rdf:Description</pre>
                            rdf:about="http://www.ePolitix.com/Authors/CraigHoy">
                             <rdf:type rdf:resource="http://www.schemas.org/Schemas#Person"/>
                             <dc:Format>text/html</dc:Format>
                             <p:Name>Craig Hoy</p:Name>
                             <p:Email>Craig.Hoy@ePolitix.com</p:Email>
                                                <rdf:RDF
                           </rdf:Description>
                                                 xmlns:sch="http://www.schemas.org/Schemas#">
                         </rdf:RDF>
                                                   rdf:about="http://www.ePolitix.com/Authors/CraigHoy">
  Permite interpretar cualquier
                                                    <dc:Format>text/html</dc:Format>
 documento XML como un
                                                    <p:Name>Craig Hoy</p:Name>
                                                    <p:Email>Craig.Hoy@ePolitix.com</p:Email>
 conjunto de tríos.
                                                  </sch:Person>
                                                </rdf:RDF>
```

Contenedores: rdf:Bag, rdf:Seq, rdf:Alt

Se usan en cualquier sitio donde pueda ir un elemento <rdf:Description>.



¿Por qué necesitamos RDF? ¿Por qué no usar sólo XML?

- El modelo de RDF es muy sencillo, se corresponde con un modelo sin ambigüedad (en XML el anidamiento de elementos no tiene un significado estándar).
- El orden de las sentencias RDF no importa (en XML el orden de los elementos es relevante, mantener el orden de millones de elementos es caro y difícil).
- Las sentencias RDF son adecuadas para tratarlas a gran escala (las posibles estructuras de XML son costosas de tratar a ese nivel).
 Oescription>The value of this property contains some

text, mixed up with child properties such as its temperature (<Temp>48</Temp>) and longitude (<Longt>101</Longt>).

</Description>

Consultas en SPARQL

¿Es necesario SPARQL?

- SPARQL es un lenguaje de consulta para RDF.
 - W3C Recommendation 15 January 2008
- ¿Necesitamos un nuevo lenguaje o podríamos usar un lenguaje de consulta para XML?
 - Un documento RDF se puede expresar en XML.
 - El mismo documento RDF se puede expresar de varias formas en XML.
 - Una consulta sólo sería válida para algunas formas de expresión (no funcionaría en otras).
- XML está a nivel de abstracción más bajo que RDF.

Select-From-Where

- SELECT: Proyección, número y orden de los datos que se obtienen.
- FROM: Fuente que se consulta, es opcional, por defecto es la base de conocimiento del sistema.
- WHERE: Restricciones sobre las posibles soluciones usando las operaciones boolenas y patrones de grafo.

31

Consultas Básicas

- Una consulta consiste en definir patrones de grafo (AND, OR de patrones) que se aplican sobre el grafo RDF.
- El resultado de las consultas son conjuntos o grafos RDF.

Ejemplos de consultas

```
SELECT ?x ?y
                                        SELECT ?x ?y
WHERE
                                        WHERE
   ?x rdf:type uni:Lecturer;
                                           ?x rdf:type uni:Lecturer.
      uni:phone ?y.
                                           ?x uni:phone?y.
}
                                        }
 SELECT ?n
                                        SELECT ?name ?email
 WHERE
                                        WHERE
                                       {
    ?x rdf:type uni:Course;
                                           ?x rdf:type uni:Lecturer;
       uni:isTaugtBy:949352.
                                              uni:name?name.
    ?c uni:name?n.
                                           OPTIONAL {?x uni:email ?email}.
    FILTER (?c = ?x).
                                                      OUTER Join
 }
```

33

Conclusiones y Bibliografía

- RDF proporciona una base para representar y procesar metadatos.
- El modelo de datos de RDF está basado en Grafos.
- RDF tiene una sintaxis en XML para soportar interoperabilidad sintáctica.
 - RDF y XML se complementan ya que RDF soporta interoperabilidad semántica.
- La filosofía de RDF es descentralizada, soporta la construcción incremental de conocimiento.

AAA Slogan: "Anyone can say Anything about Any topic."

- RDF es independiente del dominio.
- Hay lenguajes de consulta para RDF, el lenguaje estándar es SPARQL.

Bibliografía

- D. Dodds, et al.: Professional XML Meta Data.Wrox Press (July 2001).
- Dublin Core Metadata Element Set, Version 1.1.
 http://dublincore.org/documents/dces/
- Uniform Resource Identifier (URI): Generic Syntax.
 http://labs.apache.org/webarch/uri/rfc/rfc3986.html
- RDF Primer. http://www.w3.org/TR/REC-rdf-syntax/
- RDF/XML Syntax Specification (Revised). http://www.w3.org/TR/rdf-syntax-grammar/
- RDFa in XHTML: Syntax and Processing. A collection of attributes and processing rules for extending XHTML to support RDF. http://www.w3.org/TR/rdfa-syntax/
- RDFa Primer. Bridging the Human and Data Webs. http://www.w3.org/TR/xhtml-rdfa-primer/
- SPARQL Query Language for RDF. http://www.w3.org/TR/rdf-sparql-query/

