

XML y XML Schema



SGML

Standard Generalized Markup Language

HTML

XML

HTML Ejemplo:

```
<h2> A Semantic Web Primer </h2>  
<i> por <b> Gregoris Antoniou </b>  
and <b> Frank van Harmelen </b>  
</i>  
<br> The MIT Press 2004 <br>
```

HTML vs XML

<h2> A Semantic Web Primer </h2>

<i> por Gregoris Antoniou and Frank van Harmelen </i>

 The MIT Press 2004
 ISBN 0-262-01210-3

<libro>

<titulo> A Semantic Web Primer </titulo>

<autor> Gregoris Antoniou </autor>

<autor> Frank van Harmelen </autor>

<editor> The MIT Press </editor>

<año> 2004 </año>

<ISBN> 0-262-01210-3 </ISBN>

</libro>

Ambos usan tags

¿Para qué sirven las tags ?

HTML vs XML

`<h2>` A Semantic Web Primer `</h2>`

`<i>` por `` Gregoris Antoniou `` and `` Frank van Harmelen

`
` The MIT Press 2004 `
` ISBN 0-262-01210-3

Para mostrar
información

`<libro>`

`<titulo>` A Semantic Web Primer `</titulo>`

`<autor>` Gregoris Antoniou `</autor>`

`<autor>` Frank van Harmelen `</autor>`

`<editor>` The MIT Press `</editor>`

`<año>` 2004 `</año>`

`<ISBN>` 0-262-01210-3 `</ISBN>`

`</libro>`

Para describir el
contenido

Ambos usan tags

¿Para qué sirven las tags ?

XML

<libro>

<titulo> A Semantic Web Primer </titulo>

<autor> Gregoris Antoniou </autor>

<autor> Frank van Harmelen </autor>

<editor> The MIT Press </editor>

<año> 2004 </año>

<ISBN> 0-262-01210-3 </ISBN>

</libro>

recupere el título del libro:

Del documento XML puede hacerlo fácilmente
¿por qué?

XML

<libro>

<titulo> A Semantic Web Primer </titulo>

<autor> Gregoris Antoniou </autor>

<autor> Frank van Harmelen </autor>

<editor> The MIT Press </editor>

<año> 2004 </año>

<ISBN> 0-262-01210-3 </ISBN>

</libro>

recupere el título del libro:

Del documento XML puede hacerlo fácilmente
¿por qué?

Porque las tags describen la información

**Debe haber acuerdo en el vocabulario
utilizado**

XML

Comunidades de diferentes dominios han definido vocabularios especializados o “extensiones”:

- NewsML define un lenguaje XML para expresar noticias. <https://iptc.org/standards/newsml-g2/>
- HR-XML Define un lenguaje XML para Recursos Humanos <https://schemas.liquid-technologies.com/HR-XML/2007-04-15/>
- XBRL(eXtensible Business Reporting Language) diseñado para intercambiar información de negocios y financiera. <https://www.xbrl.org/thestandard/what/an-introduction-to-xbrl/>
- MathML lenguaje XML para expresar contenido científico y matemático en la web. <https://www.w3.org/Math/whatIsMathML.html>
- phyloXML es un lenguaje XML para biología evolutiva y genómica comparativa <http://www.phyloxml.org/> • XML Metadata Interchange Specification XML para intercambio de documentos. <https://www.omg.org/spec/XMI/About-XMI/>
- SBML: The Systems Biology Markup Language lenguaje XML para representar procesos biológicos. <http://co.mbine.org/standards/sbml>
- Augmented Reality Markup Language (ARML) lenguaje XML para describir escenas de realidad aumentada. <https://www.ogc.org/standards/arml>
- Travel <https://schemas.liquid-technologies.com/OpenTravel/2008B/>

XML

Un documento XML contiene:

- Declaración y referencia a documento de estructura (opcional)

```
<?xml version = "1.0" encoding = "UTF-16" ?>
```

- Elementos: contenido descrito por el documento
 - Definiciones de etiquetas (tag)

Elemento: contiene elementos y datos

Dentro de un elemento: el orden e los sub-elementos y los datos es relevante

XML

ELEMENTOS se pueden anidar en profundidad arbitraria y pueden tener propiedades

```
< persona >
```

```
  < nombre idioma="español" > Juan </ nombre >
```

```
  < edad > 32 </ edad >
```

```
  < e-mail > jjj@abc.com </ e-mail >
```

```
< /persona >
```

```
< persona >
```

```
  < nombre idioma="inglés" > John </ nombre >
```

```
  < edad > 32 </ edad >
```

```
  < e-mail > jjj@abc.com </ e-mail >
```

```
< /persona >
```

XML

Documentos bien formados: Documento Válido

- Etiquetas balanceadas → **Sintaxis**

< persona >

 < nombre idioma="español" > Juan </nombre >

 < edad > 32

 < e-mail > jjj@abc.com < /e-mail >

< /persona >

XML

¿Un documento XML validado es suficiente para intercambio de documentos?

XML

¿Un documento XML validado es suficiente para intercambio de documentos?

Poca consistencia: las etiquetas las entiende quien construye el document, son arbitrarias.

XML

¿Es suficiente para intercambio de documentos?

- **Poca consistencia:** las etiquetas las entiende quien construye el documento, son arbitrarias.

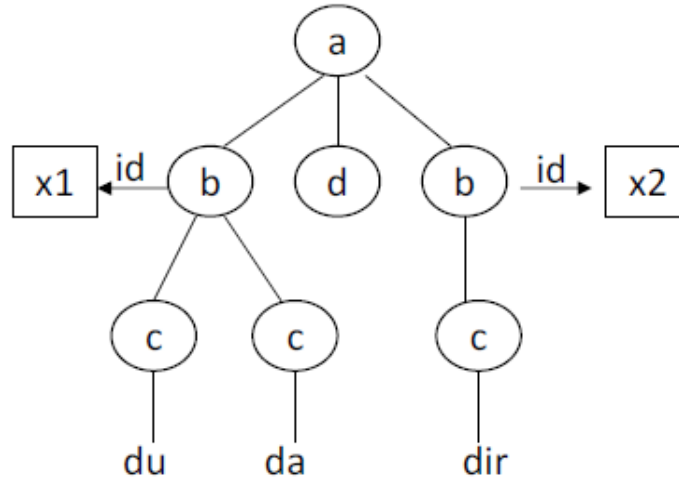
Elemento: contiene elementos y datos

Dentro de un elemento: el orden e los sub-elementos y los datos es relevante

Datos contenidos en un árbol rotulado

Documentos XML y Esquemas

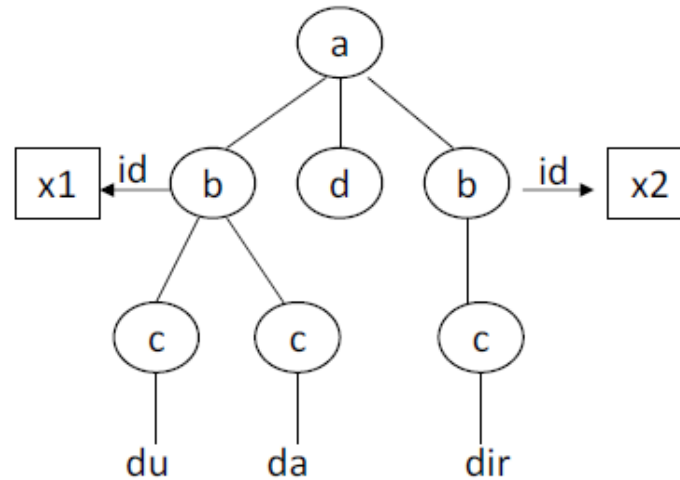
```
< ?XML version="1.0"? >  
< a >  
  < b id="x1" >  
    < c > du < /c >  
    < c > du < /c >  
  < /b >  
< d />  
  < b id="x2" >  
    < c > dir < /c >  
  < /b >  
< /a >
```



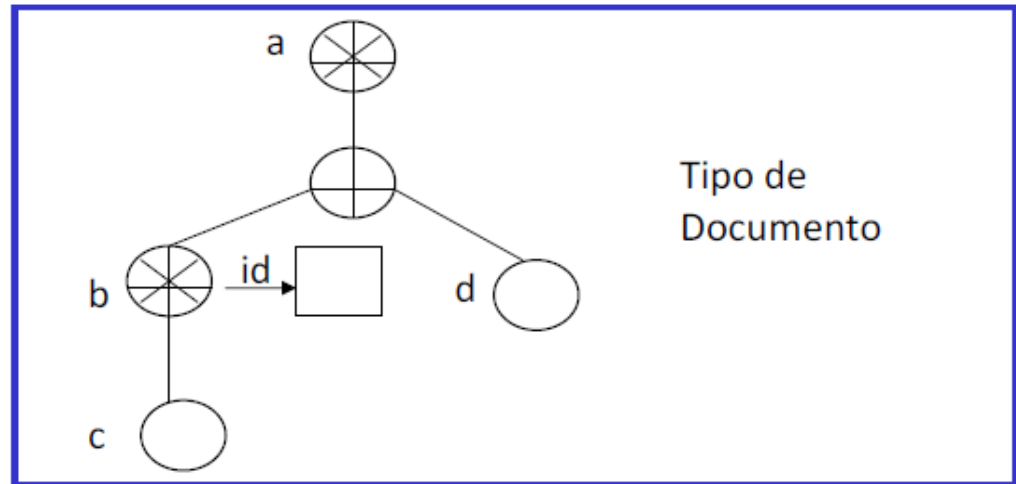
Estructura
Lógica

Documentos XML y Esquemas

```
< ?XML version="1.0"? >  
< a >  
  < b id="x1" >  
    < c > du < /c >  
    < c > du < /c >  
  < /b >  
  < d />  
  < b id="x2" >  
    < c > dir < /c >  
  < /b >  
< /a >
```



Estructura
Lógica



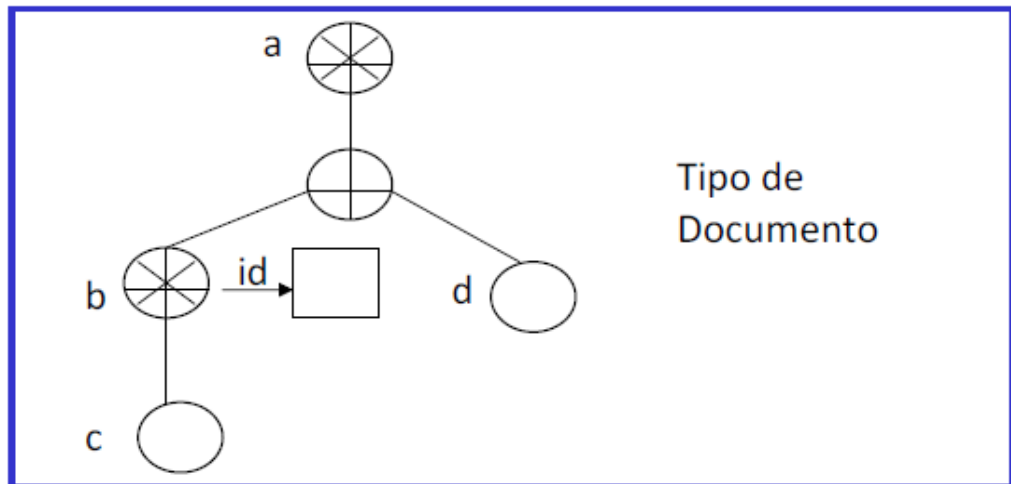
Tipo de
Documento

XML Schema

- Describe si un documento XML posee cierta forma

Tipo de documento

EJEMPLO:



Documentos XML - Orden

Documentos estructurados:

Persona: {nombre: “Juan”, apellido : “Perez”}

Persona: {apellido : “Perez”, nombre: “Juan”}

Equivalentes

Documentos semi-estructurados XML:

```
< persona >  
  < nombre > Juan < /nombre >  
  < apellido > Perez < /apellido >  
< /persona >
```

```
< persona >  
  < apellido > Perez < /apellido >  
  < nombre > Juan < /nombre >  
< /persona >
```

NO Equivalentes

Documento XML

- `<bookstore xmlns="http://www.contoso.com/books">`
- `<book genre="autobiography" publicationdate="1981-03-22" ISBN="1-861003-11-0">`
- `<title>The Autobiography of Benjamin Franklin</title>`
- `<author>`
- `<first-name>Benjamin</first-name>`
- `<last-name>Franklin</last-name>`
- `</author>`
- `<price>8.99</price>`
- `</book>`
- `<book genre="novel" publicationdate="1967-11-17" ISBN="0-201-63361-2">`
- `<title>The Confidence Man</title>`
- `<author>`
- `<first-name>Herman</first-name>`
- `<last-name>Melville</last-name>`
- `</author>`
- `<price>11.99</price>`
- `</book>`
- `<book genre="philosophy" publicationdate="1991-02-15" ISBN="1-861001-57-6">`
- `<title>The Gorgias</title>`
- `<author>`
- `<name>Plato</name>`
- `</author>`
- `<price>9.99</price>`
- `</book>`
- `</bookstore>`

XML Schema

- `<?xml version="1.0" encoding="utf-8"?>`
- `<xs:schema attributeFormDefault="unqualified" elementFormDefault="qualified"`
- `targetNamespace="http://www.contoso.com/books" xmlns:xs="http://www.w3.org/2001/XMLSchema">`
- `<xs:element name="bookstore">`
- `<xs:complexType>`
- `<xs:sequence>`
- `<xs:element maxOccurs="unbounded" name="book">`
- `<xs:complexType>`
- `<xs:sequence>`
- `<xs:element name="title" type="xs:string" />`
- `<xs:element name="author">`
- `<xs:complexType>`
- `<xs:sequence>`
- `<xs:element minOccurs="0" name="name" type="xs:string" />`
- `<xs:element minOccurs="0" name="first-name" type="xs:string" />`
- `<xs:element minOccurs="0" name="last-name" type="xs:string" />`
- `</xs:sequence>`
- `</xs:complexType>`
- `</xs:element>`
- `<xs:element name="price" type="xs:decimal" />`
- `</xs:sequence>`
- `<xs:attribute name="genre" type="xs:string" use="required" />`
- `<xs:attribute name="publicationdate" type="xs:date" use="required" />`
- `<xs:attribute name="ISBN" type="xs:string" use="required" />`
- `</xs:complexType>`
- `</xs:element>`
- `</xs:sequence>`
- `</xs:complexType>`
- `</xs:element>`
- `</xs:schema>`

Documentos XML y Esquemas – XML Schema

```
<xsd:schema xmlns:xsd="http://www.w3.org/2001/XMLSchema">
```

XML Namespace

```
<xsd:element name="hojaPedido" type="TipoHojaPedido"/>  
<xsd:element name="comentario" type="xsd:string"/>
```

Elementos

```
<xsd:complexType name="TipoHojaPedido">  
  <xsd:sequence>  
    <xsd:element name="facturarA" type="direccionEEUU"/>  
    <xsd:element name="elementos" type="Elementos"/>  
  </xsd:sequence>  
  <xsd:attribute name="fechaPedido" type="xsd:date"/>  
</xsd:complexType>
```

Definición de tipos

```
<xsd:complexType name="direccionEEUU">  
  <xsd:sequence>  
    <xsd:element name="nombre" type="xsd:string"/>  
    <xsd:element name="calle" type="xsd:string"/>  
    <xsd:element name="ciudad" type="xsd:string"/>  
  </xsd:sequence>  
  <xsd:attribute name="pais" type="xsd:NMTOKEN" fixed="EEUU"/>  
</xsd:complexType>
```

Documentos XML y Esquemas – XML Schema

```
<xsd:complexType name="Elementos">
  <xsd:sequence>
    <xsd:element name="elemento" minOccurs="0" maxOccurs="unbounded">
      <xsd:complexType>
        <xsd:sequence>
          <xsd:element name="nombreProducto" type="xsd:string"/>
          <xsd:element name="cantidad">
            <xsd:simpleType>
              <xsd:restriction base="xsd:positiveInteger">
                <xsd:maxExclusive value="100"/>
              </xsd:restriction>
            </xsd:simpleType>
          </xsd:element>
          <xsd:element name="precioEEUU" type="xsd:decimal"/>
          <xsd:element ref="comentario" minOccurs="0"/>
          <xsd:element name="fechaEnvio" type="xsd:date" minOccurs="0"/>
        </xsd:sequence>
        <xsd:attribute name="numProducto" type="SKU" use="required"/>
      </xsd:complexType>
    </xsd:element>
  </xsd:sequence>
</xsd:complexType>
```

.....

Herramientas para validar

<https://www.liquid-technologies.com/online-xsd-validator>

Enlaces con material

<http://www.xml.com/axml/testaxml.htm>

<http://www.w3c.es/Traducciones/es/TR/2001/REC-xmlschema-0-20010502/>

<http://www.w3c.org/TR/REC-xml-names/>