

# Actividad 6

## Creación De Documento XML Y DTD Según Especificaciones



Rubén Beltrán Muñoz

Lenguaje de Marcas

1º DAM

## Índice

1. Crear un XML con el modelo de datos indicado en el enunciado. ....2
2. Elaborar un DTD que permita validar el documento XML. ....4
3. Validar el documento con alguna aplicación externa. ....5
4. Con el mismo XML utilizar el estándar XML Schema. ....7

## 1. Crear un XML con el modelo de datos indicado en el enunciado.

Para la creación del XML se empiezan a marcar los puntos que la van a componer. **Elementos raíz,** **elementos de primarios/secundarios,** **atributos**

- Existen tres tipos de documentos almacenados en la **biblioteca**: **libros, revistas y periódicos**. Todos los documentos están identificados por el atributo **Id**.
  - Para los **libros** este **atributo** empieza con la letra "L" seguido de 4 dígitos identificativos.
  - Para los **periódicos** este **atributo** empieza por la letra "P" seguido de los 4 dígitos identificativos.
  - En el caso de las **revistas** empieza por la letra "R".
- Los **libros** a su vez son clasificados en **novela, infantil o didáctico**. Cada libro contiene un **atributo** **identificativo de su clase denominado tipo\_clase**. Dentro de cada **libro** se tiene un **título, varios capítulos** con el **título en su interior, un índice y una sinopsis**. Tanto en libro como en capítulo existe un **atributo que contiene el número de páginas** del libro o del capítulo, según corresponda. Cada capítulo contiene un **elemento denominado contenido**, en el que se tiene un **atributo con el enlace a la información**.
- Las **revistas** a su vez son clasificadas en: **informática, corazón, coches, investigación y otras**. Cada revista tiene el **atributo tipo\_clase** identificativo de la clase a la que pertenece. Dentro de cada revista tenemos **el título, el número de la revista, un índice de contenido y las secciones**. En cada sección y en la **revista** se tiene un **atributo que contiene el número de páginas**. Además, en cada **sección se tiene la parte denominada contenido**, en la que se tiene un **atributo con un enlace a la información**. De igual forma que en los anteriores, la última etiqueta del árbol debe ser el **contenido con un atributo que referencia al contenido**.
- Los **periódicos** se clasifican en nacionales e internacionales. Cada **periódico** contiene el **atributo tipo\_clase** identificativo de la clase a la que pertenece y un **atributo que incluye la fecha de publicación**.
- Dentro de los **periódicos** tenemos **secciones y un índice**. Cada sección debe contener un **atributo** **identificativo del tipo de sección, que puede ser: económica, opinión, deportes, nacional o internacional**. Las **secciones se dividen en artículos**, en donde se define en un **atributo el autor**. Finalmente, **el contenido será el último elemento** del árbol, que necesita un atributo que referencia a la información.

### ➤ **Biblioteca**

- **Libros** (Atr. **Id**, **tipo\_clase**, **num\_pag**)
  - Título
  - Capítulos (Atr. **Num-pag**)
    - Título
    - Contenido (Atr. **URL**)
  - Índice
  - Sinopsis
- **Revistas** (Atr. **Id**, **tipo\_clase**)
  - Título
  - Num\_revista
  - Índice
  - Sección (Atr. **num\_pag**)
    - Contenido (Atr. **URL**)
- **Periódicos** (Atr. **Id**, **tipo\_clase**, **fecha\_publicacion**)
  - Sección (Atr. **tipo\_seccion**)
    - Artículo (Atr. **Autor**)
  - Índice
  - Contenido (Atr. **Url**)

```

<?xml version="1.0" encoding="UTF-8"?>
<biblioteca>
  <libros id="L001" tipo_clase="novela" num_pag="879">
    <titulo>Fuego y Sangre</titulo>
    <capitulo num_pag="28">
      <titulo>La Conquista De Aegon</titulo>
      <contenido url="Cap1.txt"/>
    </capitulo>
    <capitulo num_pag="35">
      <titulo>El Reinado del Dragón</titulo>
      <contenido url="Cap2.txt"/>
    </capitulo>
    <indice>
      1.La Conquista De Aegon
      2.El Reinado del Dragón.
      3.Las Tres Cabezas Del Dragon
    </indice>
    <sinopsis>
      Siglos antes de que tuvieran lugar los acontecimientos que se relatan en «Canción de hielo y fuego»,
      la casa Targaryen, la única dinastía de señores dragón que sobrevivió a la Maldición de Valyria,
      se asentó en la isla de Rocadragón
    </sinopsis>
  </libros>

  <revistas id="R015" tipo_clase="informatica">
    <titulo>Mi Primera Web</titulo>
    <num_revista>128</num_revista>
    <indice>
      1.Etiquetas
      2.HTML
      3.CSS
    </indice>
    <seccion num_pag="12">
      <contenido url="Etiquetas.txt"/>
    </seccion>
  </revistas>

  <periodicos id="P157" tipo_clase="nacionales" fecha_publicacion="21/01/2019">
    <seccion tipo_seccion="deportes">
      <articulo autor="Michael Robison">
        Mucho se discute, en España y fuera, sobre si el Barcelona de Pep Guardiola es el mejor
        equipo de fútbol de todos los tiempos. Por más entretenido que nos resulte, es un debate
        imposible de resolver. Es incluso estéril. Existen demasiadas variantes y nos falta información.
      </articulo>
    </seccion>
    <indice>
      1. ¿Habra Triplete?
      2. Detras del VAR
      3. El Barcelona de Pep
    </indice>
    <contenido url="MRP157.txt"/>
  </periodicos>
</biblioteca>

```

## 2. Elaborar un DTD que permita validar el documento XML.

```
<!DOCTYPE biblioteca [  
<!ELEMENT biblioteca ( libros | revistas | periodicos)+>  
  
<!ELEMENT libros (titulo, (capitulo+), indice, sinopsis)>  
<!ELEMENT titulo (#PCDATA)>  
<!ELEMENT capitulo (titulo,contenido)>  
<!ELEMENT contenido (#PCDATA)>  
<!ELEMENT indice (#PCDATA)>  
<!ELEMENT sinopsis (#PCDATA)>  
  
<!ELEMENT revistas (titulo,num_revista,indice,seccion+)>  
<!ELEMENT num_revista (#PCDATA)>  
<!ELEMENT seccion (contenido | (articulo+))>  
  
<!ELEMENT periodicos ((seccion+), indice, contenido)>  
<!ELEMENT articulo (#PCDATA)>  
  
<!ATTLIST libros  
    id ID #REQUIRED  
    tipo_clase ( novela | infantil | didactico) "infantil"  
    num_pag CDATA #REQUIRED>  
  
<!ATTLIST capitulo num_pag CDATA #REQUIRED>  
  
<!ATTLIST contenido url CDATA #IMPLIED>  
  
<!ATTLIST revistas  
    id ID #REQUIRED  
    tipo_clase ( informatica | corazon | coches | investigacion | otras ) "coches">  
  
<!ATTLIST seccion num_pag CDATA #IMPLIED>  
  
<!ATTLIST periodicos  
    id ID #REQUIRED  
    tipo_clase ( nacionales | internacionales) "nacionales"  
    fecha_publicacion CDATA #REQUIRED >  
  
<!ATTLIST seccion tipo_seccion (economica | opinion | deportes | nacional | internacional) "opinion">  
<!ATTLIST articulo autor CDATA #REQUIRED>  
>|
```

### 3. Validar el documento con alguna aplicación externa.

- Para este punto utilizo <http://www.xmlvalidation.com>
- Para este punto he unido la parte de XML y DTD en un solo bloc de notas para validarlo.

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE biblioteca [
<ELEMENT biblioteca ( libros | revistas | periodicos)+>

<ELEMENT libros (titulo, (capitulo+), indice, sinopsis)>
<ELEMENT titulo (#PCDATA)>
<ELEMENT capitulo (titulo,contenido)>
<ELEMENT contenido (#PCDATA)>
<ELEMENT indice (#PCDATA)>
<ELEMENT sinopsis (#PCDATA)>

<ELEMENT revistas (titulo,num_revista,indice,seccion+)>
<ELEMENT num_revista (#PCDATA)>
<ELEMENT seccion (contenido | (articulo+))>

<ELEMENT periodicos ((seccion+), indice, contenido)>
<ELEMENT articulo (#PCDATA)>

<!ATTLIST libros
    id ID #REQUIRED
    tipo_clase ( novela | infantil | didactico) "infantil"
    num_pag CDATA #REQUIRED>

<!ATTLIST capitulo num_pag CDATA #REQUIRED>
<!ATTLIST contenido url CDATA #IMPLIED>

<!ATTLIST revistas
    id ID #REQUIRED
    tipo_clase ( informatica | corazon | coches | investigacion | otras ) "coches">

<!ATTLIST seccion num_pag CDATA #IMPLIED>

<!ATTLIST periodicos
    id ID #REQUIRED
    tipo_clase ( nacionales | internacionales) "nacionales"
    fecha_publicacion CDATA #REQUIRED >

<!ATTLIST seccion tipo_seccion (economica | opinion | deportes | nacional | internacional) "opinion">
<!ATTLIST articulo autor CDATA #REQUIRED>
]>

<biblioteca>
    <libros id="L001" tipo_clase="novela" num_pag="879">
        <titulo>Fuego y Sangre</titulo>
        <capitulo num_pag="28">
            <titulo>La Conquista De Aegon</titulo>
            <contenido url="Cap1.txt"/>
        </capitulo>
        <capitulo num_pag="35">
            <titulo>El Reinado del Dragón</titulo>
            <contenido url="Cap2.txt"/>
        </capitulo>
    </libros>

```

```

<indice>
  1.La Conquista De Aegon
  2.El Reinado del Dragón.
  3.Las Tres Cabezas Del Dragon
</indice>
<sinopsis>
  Siglos antes de que tuvieran lugar los acontecimientos que se relatan en «Canción de hielo y fuego»,
  la casa Targaryen, la única dinastía de señores dragón que sobrevivió a la Maldición de Valyria,
  se asentó en la isla de Rocradragón
</sinopsis>
</libros>

<revistas id="R015" tipo_clase="informatica">
  <titulo>Mi Primera Web</titulo>
  <num_revista>128</num_revista>
  <indice>
    1.Etiquetas
    2.HTML
    3.CSS
  </indice>
  <seccion num_pag="12">
    <contenido url="Etiquetas.txt"/>
  </seccion>
</revistas>

<periodicos id="P157" tipo_clase="nacionales" fecha_publicacion="21/01/2019">
  <seccion tipo_seccion="deportes">
    <articulo autor="Michael Robison">
      Mucho se discute, en España y fuera, sobre si el Barcelona de Pep Guardiola es el mejor
      equipo de fútbol de todos los tiempos. Por más entretenido que nos resulte, es un debate
      imposible de resolver. Es incluso estéril. Existen demasiadas variantes y nos falta información.
    </articulo>
  </seccion>
  <indice>
    1. ¿Habra Triplete?
    2. Detras del VAR
    3. El Barcelona de Pep
  </indice>
  <contenido url="MRP157.txt"/>
</periodicos>
</biblioteca>

```

No me permite declarar el elemento sección nuevamente, con lo que implemento artículo en la declaración de sección declarada para revistas.

#### Errors in the XML document:

❌ 17: 30 Element type "seccion" must not be declared more than once.

```

<!ELEMENT num_revista (#PCDATA)>
<!ELEMENT seccion (contenido | articulo)>
<!ELEMENT periodicos (seccion indice, contenido)>
<!ELEMENT seccion (articulo)>
<!ELEMENT articulo (#PCDATA)>

```

EL documento se encuentra sin errores, he tenido que corregir algunos errores gramaticales (periódico, en vez de periódicos; signosis, escrito en la declaración DTD como sinopsis...) que me estaban causando problemas, pero gracias al validador los detectaba correctamente

No errors were found



#### 4. Con el mismo XML utilizar el estándar XML Schema.

- Para este punto utilizo la web <https://www.freeformatter.com/xsd-generator.html#ad-output> Solo tengo que copiar mi archivo XML completo con la parte de DTD incluida.

Option 1: Copy-paste your XML document here

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE biblioteca [
<!ELEMENT biblioteca ( libros | revistas | periodicos)+>

<!ELEMENT libros (titulo,capitulo+, indice, sinopsis)>
<!ELEMENT titulo (#PCDATA)>
```

Option 2: Or upload your XML document

Seleccionar archivo Ningún archivo seleccionado UTF-8

XSD Design:

Russian Doll

GENERATE XSD GENERATE XSD IN NEW WINDOW

1 | Free SQL Course Learn SQL for Data Science in 4 hours by completing 41 interactive exercises. DataCamp

2 | Get Postman Enterprise API platform optimized for professional dev teams. 100K+ companies worldwide use Postman Postman

Generated XSD:

```
<xs:complexType>
  <xs:sequence>
    <xs:element type="xs:string" name="titulo"/>
    <xs:element name="contenido">
      <xs:complexType>
        <xs:simpleContent>
          <xs:extension base="xs:string">
            <xs:attribute type="xs:string" name="url" use="optional"/>
          </xs:extension>
        </xs:simpleContent>
      </xs:complexType>
    </xs:element>
  </xs:sequence>
  <xs:attribute type="xs:byte" name="num_pag" use="optional"/>
</xs:complexType>
</xs:element>
<xs:element type="xs:string" name="indice"/>
<xs:element type="xs:string" name="sinopsis"/>
</xs:sequence>
<xs:attribute type="xs:string" name="id"/>
```

COPY TO CLIPBOARD SELECT ALL



## Parte de XML SCHEMA

```
<xs:schema attributeFormDefault="unqualified" elementFormDefault="qualified" xmlns:xs="http://www.w3.org/2001/XMLSchema">
  <xs:element name="biblioteca">
    <xs:complexType>
      <xs:sequence>
        <xs:element name="libros">
          <xs:complexType>
            <xs:sequence>
              <xs:element type="xs:string" name="titulo"/>
              <xs:element name="capitulo" maxOccurs="unbounded" minOccurs="0">
                <xs:complexType>
                  <xs:sequence>
                    <xs:element type="xs:string" name="titulo"/>
                    <xs:element name="contenido">
                      <xs:complexType>
                        <xs:simpleContent>
                          <xs:extension base="xs:string">
                            <xs:attribute type="xs:string" name="url" use="optional"/>
                          </xs:extension>
                        </xs:simpleContent>
                      </xs:complexType>
                    </xs:element>
                  </xs:sequence>
                <xs:attribute type="xs:byte" name="num_pag" use="optional"/>
              </xs:complexType>
            </xs:element>
            <xs:element type="xs:string" name="indice"/>
            <xs:element type="xs:string" name="sinopsis"/>
          </xs:sequence>
          <xs:attribute type="xs:string" name="id"/>
          <xs:attribute type="xs:string" name="tipo_clase"/>
          <xs:attribute type="xs:short" name="num_pag"/>
        </xs:complexType>
      </xs:element>
      <xs:element name="revistas">
        <xs:complexType>
          <xs:sequence>
            <xs:element type="xs:string" name="titulo"/>
            <xs:element type="xs:short" name="num_revista"/>
            <xs:element type="xs:string" name="indice"/>
            <xs:element name="seccion">
              <xs:complexType>
                <xs:sequence>
                  <xs:element name="contenido">
                    <xs:complexType>
                      <xs:simpleContent>
                        <xs:extension base="xs:string">
                          <xs:attribute type="xs:string" name="url"/>
                        </xs:extension>
                      </xs:simpleContent>
                    </xs:complexType>
                  </xs:element>
                </xs:sequence>
                <xs:attribute type="xs:byte" name="num_pag"/>
                <xs:attribute type="xs:string" name="tipo_seccion"/>
              </xs:complexType>
            </xs:element>
          </xs:sequence>
          <xs:attribute type="xs:string" name="id"/>
          <xs:attribute type="xs:string" name="tipo_clase"/>
        </xs:complexType>
      </xs:element>
      <xs:element name="periodicos">
        <xs:complexType>
          <xs:sequence>
            <xs:element name="seccion">
              <xs:complexType>
                <xs:sequence>
                  <xs:element name="articulo">
                    <xs:complexType>
                      <xs:simpleContent>
                        <xs:extension base="xs:string">
                          <xs:attribute type="xs:string" name="autor"/>
                        </xs:extension>
                      </xs:simpleContent>
                    </xs:complexType>
                  </xs:element>
                </xs:sequence>
                <xs:attribute type="xs:string" name="tipo_seccion"/>
              </xs:complexType>
            </xs:element>
          </xs:sequence>
          <xs:attribute type="xs:string" name="tipo_seccion"/>
        </xs:complexType>
      </xs:element>
    </xs:sequence>
  </xs:element>
</xs:schema>
```