

Ciclos Formativos de Grado Superior

Lenguajes de Marcas y Sistemas de Gestión de la Información

Capítulo 4: Definición de esquemas y vocabularios en XML



Javier S. Zurdo, Pablo Toharia, Laura Raya



Índice

1. DTD
2. Esquemas
3. Validación de documentos XML



DTD

- **DTD** significa «*Document Type Definition*»
- Define cómo se construye un documento XML
- Establece:
 - Qué elementos son aceptados
 - En qué posiciones deben estar dentro de un documento XML.

¿Por qué resulta importante la creación de DTD?

- Porque es bueno compartir
- Porque debe estar sujeto a una serie de reglas para que todos los que lo utilicen de igual manera.
 - Consultas, inserciones, modificaciones, borrados, ...

DTD

- **Ejemplo, almacenar los siguientes datos de un sms:**

- Número de teléfono del usuario
- Fecha de envío
- Hora de envío
- Contenido del mensaje

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<!DOCTYPE BDsms SYSTEM "BDsms.dtd">
<BDsms>
  <sms>
    <telefono>955 55 66 55</telefono>
    <fecha>1/7/2011</fecha>
    <hora>23:55</hora>
    <mensaje>Juego1: Tetris</mensaje>
  </sms>
  ...
</BDsms>
```

DTD

- **Ejemplo, almacenar los siguientes datos de un sms:**

- La estructura (en el fichero BDsms.dtd):

```
<!ELEMENT BDsms (sms*)>  
<!ELEMENT sms (telefono, fecha, hora, mensaje)>  
<!ELEMENT telefono (#PCDATA)>  
<!ELEMENT fecha (#PCDATA)>  
<!ELEMENT hora (#PCDATA)>  
<!ELEMENT mensaje (#PCDATA)>
```

DTD

- **Bloques para construir una DTD:**
 - Elemento: por ejemplo teléfono
 - Atributo: Para añadir más información a un elemento
 - Entidad: ‘ ‘ < > “ ’ &
 - PCDATA: Parsed Character Data
 - CDATA: Character Data
- **Secuencias de elementos: Estructura con hijos**
 - ‘+’ 1 ó + ocurrencias
 - ‘*’ 0 ó + ocurrencias
 - ‘?’ 0 ó 1 ocurrencia

DTD

- **Realizar la siguiente actividad:**

Se quiere crear una biblioteca de libros en un documento XML. Para ello se necesita crear una DTD que almacene los siguientes campos de un libro:

- Código de libro
- Título
- Editorial
- Edición
- ISBN
- Número de páginas
- Autor

Esquemas

- XML Schema , también llamado XSD o «XML Schema Definition»
- Permite generar un lenguaje XML con la definición de las reglas que lo componen (como los DTD)
- Características:
 - Define qué elementos pueden aparecer en un documento XML
 - Define qué atributos pueden aparecer en un documento XML
 - Define qué elementos son compuestos, indicando qué elementos hijos deben aparecer y en qué orden
 - Define qué elementos pueden ser vacíos o que pueden incluir texto asociado
 - Define los tipos que pueden utilizarse en cada elemento o atributo
 - Define la obligatoriedad, la optatividad de elementos y/o atributos

Esquemas

- XML Schema
- Ventajas del Esquema vs DTD:
 - Estándar desde el 2001 de W3C
 - Extensible, más potente y gramáticas más complejas
 - Definición de tipos de datos

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<BDsms xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.gmr.v.es BDsmsSchema.xsd">
  <sms>
    <telefono>955 55 66 55</telefono>
    <fecha>1/7/2011</fecha>
    <hora>23:55</hora>
    <mensaje>Juego1: Tetris</mensaje>
  </sms>
  ...
</BDsms>
```

Esquemas

- XML Schema, ejemplo BDsmsSchema.xsd :

```
<?xml version="1.0" ?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema" version="0.1" xml:lang="es">
  <xs:element name="BDsms">
    <xs:complexType>
      <xs:sequence>
        <xs:element name="sms" maxOccurs="unbounded">
          <xs:complexType>
            <xs:sequence>
              <xs:element name="telefono" type="xs:string"/>
              <xs:element name="fecha" type="xs:string"/>
              <xs:element name="hora" type="xs:string"/>
              <xs:element name="mensaje" type="xs:string"/>
            </xs:sequence>
          </xs:complexType>
        </xs:element>
      </xs:sequence>
    </xs:complexType>
  </xs:element>
</xs:schema>
```

Javier S. Zurdo, Pablo Toharia, Laura Raya

Esquemas

- **Bloques para construir un Esquema:**
 - Elemento raíz: `<xs:schema> ... </xs:schema>`
 - Elemento simple: `<xs:element name="nombre_elem" type="tipo_elem">`
 - `xs:string`
 - `xs:date`
 - `xs:integer`
 - Atributo: `<curso letra="A">1</curso>`
 - Restricción: `<xs:restriction base="xs:integer"> <xs:minInclusive value="16"/><xs:maxInclusive value="24"/> </xs:restriction>`
 - Elementos complejos
 - Secuencia de elementos. Indicadores:
 - `<xs:all>`, `<xs:Choice>`, `<xs:sequence>`
 - `<xs:maxOccurs>`, `<xs:minOccurs>`

Validación de documentos XML


- La validación de los documentos XML es una serie de comprobaciones que permiten saber si el documento XML está bien formado y si se ajusta a una estructura previamente definida (ya sea DTD o esquemas)
 - Cuando se habla de documento bien formado se quiere decir que se siguen las normas y reglas básicas que se establecen en todos los documentos XML.
 - Cuando se habla de un documento válido, quiere decir que además de estar bien formado, su estructura es acorde con las normas que se dictan en la definición del tipo de documento o esquema asociado.

Validación de documentos XML

- Se necesita validar los documentos para:
 - Asegurar que en una transferencia de información XML entre un emisor y un receptor, ambos interpretarán su contenido de igual manera.
 - Asegurar que el documento contiene toda la información declarada como obligatoria.
 - Asegurar que los datos vendrán en un formato conocido y correcto.
- Documento bien formado:
 - Al principio del documento se debe declarar una cabecera (en formato etiqueta) que indique lo siguiente:
 - Versión del documento XML
 - Tipo de codificación utilizada
 - Documento de estructura externo (DTD o esquema)

Validación de documentos XML

- Documento bien formado (continuación):
 - Todas las entidades, elementos y atributos deben tener una sintaxis correcta:
 - Distinción entre mayúsculas y minúsculas
 - Nombres pueden ser alfanuméricos (empezando por letra)
 - Elementos abiertos deben cerrarse. `<etiq> </etiq>`.
 - Valor de atributo entre comillas dobles o simples. `<etiq atributo="1">`
 - Elemento vacío EMPTY, debe autocerrarse. `<etiq/>`
- Documento válido (bien formado y acorde a un DTD o esquema)
 - Usar solo elementos o atributos definidos en el DTD o esquema
 - Elementos y atributos ordenados según DTD o esquema
 - Valores declarados como únicos, deben cumplirse en todo el documento
 - Declaración de tipos fuerte


Markup Validation Service
 Check the markup (HTML, XHTML, ...) of Web documents

☒ Validate by URI
 ☐ Validate by File Upload
 ☐ Validate by Direct Input

Validate by direct input


Enter the Markup to validate:

```
<?xml version="1.0" encoding="UTF-8" ?>
<!DOCTYPE message SYSTEM "OPFDATA">
<?>

<@lang>
  <msg>
    <telFonno>955 66 55</telFonno>
    <echos>17/2011</echos>
    <chomo>15-06</chomo>
    <mensaje>Juegol: Tetris</mensaje>
  </msg>
  <msg>
    <telFonno>745 15 56 11</telFonno>
    <echos>27/2011</echos>
    <chomo>15-06</chomo>
    <mensaje>Juegol: Arkxoid</mensaje>
  </msg>
  <msg>
    <telFonno>842 35 22 00</telFonno>
    <echos>28/11/2011</echos>
    <chomo>09-22</chomo>
    <mensaje>Juegol: Comecocos</mensaje>
  </msg>
</@lang>
```

More Options

Check



[Documentation](#)
[An example](#)
[About](#)

XHTML Validation

Read here how to validate your XML files (including referenced DTDs) online with just a few mouse clicks.

Please copy your XML document in here:

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<!DOCTYPE BIDsml [
  <!--
  <ELEMENT BIDsml (msms*)
  <ELEMENT sms (bidsfms, fecha, hora, mensaje)
  <ELEMENT telefono (RFCDATA*)
  <ELEMENT fecha (RFCDATA*)
  <ELEMENT hora (RFCDATA*)
  <ELEMENT mensaje (RFCDATA*)
  -->
]>
```

Or upload it:

No se ha s...n archivado

The validation check is performed against any XML schema or DTD declared inside the XML document. If neither an XML schema nor a DTD is declared, only a syntax check is performed.

To validate the XML document against an external XML schema, click below.

☐ Validate against external XML schema

XML Validator

[Home](#) | [Forum](#) | [Resources](#) | [Sidebar](#) | [Contact](#) | [About](#) | [Weblog](#)

Source

URL:

Upload
 No se ha seleccionado ningun archivo

☒ View Sourcecode ☒ Well-Formedness only

Source code

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<DTD? ID?xmlns
<ELEMENT ID?xmlns (xml*)
<ELEMENT sms (ID?xmlns, fecha, hora, mensaje)
<ELEMENT telefono (ID?DATA)?
<?ELEMENT fecha (ID?DATA)?
```

The document is well-formed

Validated files

	File name	Charsets	Source
1	SOURCECODE	ISO-8859-1	XML declaration