

Lenguajes de marcas y sistemas de gestión de la información

Declaración de tipo de documento (DTD)

Qué es DTD

DTD (Document Type Definition, Definición de Tipo de Documento) sirve para definir la estructura de un documento XML, permitiendo su validación.

Un documento XML es válido (valid) cuando, además de estar bien formado, **cumple con las normas establecidas en su estructura, es decir, cumple lo indicado en la DTD a la que esté asociado.**

De este modo grupos independientes de personas pueden acordar una DTD estándar para intercambiar datos y una aplicación puede usar esta DTD para verificar que sus datos XML son válidos.

Existen otros métodos que también permiten validar documentos XML, como por ejemplo XML Schema, que veremos más adelante.

Declaración de tipo de documento

Una DTD se puede escribir tanto interna como externamente a un archivo XML.

Cuando la declaración es **interna**, se declara dentro del archivo XML:

```

<?xml version="1.0"?>
<!DOCTYPE note [
<!ELEMENT note (to,from,heading,body)>
<!ELEMENT to (#PCDATA)>
<!ELEMENT from (#PCDATA)>
<!ELEMENT heading (#PCDATA)>
<!ELEMENT body (#PCDATA)>
]>
<note>
  <to>Tove</to>
  <from>Jani</from>
  <heading>Reminder</heading>
  <body>Don't forget me this weekend</body>
</note>

```

El DTD anterior se interpreta así:

! DOCTYPE note define que el elemento raíz de este documento es note

! ELEMENT note define que el elemento note debe contener cuatro elementos: "to, from, header, body"

! ELEMENT to define el elemento to como del tipo "#PCDATA"

! ELEMENT from define el elemento from como del tipo "#PCDATA"

! ELEMENT heading define el elemento heading como del tipo "#PCDATA"

! ELEMENT body define que el elemento del cuerpo debe ser del tipo "#PCDATA"

Cuando una declaración es **externa**, la DTD se declara en un archivo externo, la definición <!DOCTYPE> debe contener una referencia al archivo DTD:

```

<?xml version="1.0"?>
<!DOCTYPE note SYSTEM "note.dtd">
<note>
  <to>Tove</to>
  <from>Jani</from>
  <heading>Reminder</heading>
  <body>Don't forget me this weekend!</body>
</note>

```

Y el archivo "note.dtd", que contiene el DTD, sería:

```
<!ELEMENT note (to,from,heading,body)>
<!ELEMENT to (#PCDATA)>
<!ELEMENT from (#PCDATA)>
<!ELEMENT heading (#PCDATA)>
<!ELEMENT body (#PCDATA)>
```



Estructura

En una DTD se pueden declarar:

- Elementos
- Atributos
- Entidades
- Notaciones

Declaración de elementos:

<!ELEMENT nombre-del-elemento tipo-de-contenido>

El tipo de contenido puede ser:

Texto -> (#PCDATA)

Contener a otros elementos hijos -> (hijo1,hijo2,hijo3)

Los hijos se deben declarar siguiendo el orden del paréntesis.

Los hijos a su vez también pueden tener hijos.

No tener contenido (estar vacío) -> EMPTY

Cualquier tipo de contenido, incluso contenido mixto -> ANY

Cuando se quiere declarar un elemento que pueda contener cualquier contenido: texto, otros elementos o una mezcla de ambos.

Cardinalidad de los elementos:

Los elementos declarados en una DTD sobre los que no actúe ningún operador de cardinalidad, tendrán que aparecer obligatoriamente una única vez, en el documento XML asociado.

Pero también podemos definir el número de veces que puede aparecer cada elemento en el documento XML asociado utilizando los operadores de cardinalidad:

Operador	Cardinalidad	Significado
?	0-1	El elemento puede aparecer 0 o 1 vez
*	0-n	El elemento puede aparecer 0 o n veces
+	1-n	El elemento puede aparecer 1 o más veces
A B	O uno ú otro	Tiene que aparecer el elemento A ó el B.

Declaración de atributos:

<!ATTLIST nombre-del-elemento nombre-del-atributo tipo-de-atributo valor-del-atributo>

El valor del atributo puede ser:

Poner un valor por defecto -> "ValorDefecto"

El atributo es obligatorio -> #REQUIRED

El atributo es opcional -> #IMPLIED

El valor del atributo es fijo -> #FIXED value

El tipo de atributo puede ser:

Datos de tipo carácter (texto) -> CDATA

Un valor de una lista -> (e1|e2|...)

Un identificador único -> ID

El ID de otro elemento -> IDREF

Una lista de otros IDs -> IDREFS

Caracteres+ . + _ + - + : (un nombre XML válido) -> NMTOKEN

Una lista de NMTOKEN -> NMTOKENS

El nombre de una entidad -> ENTITY

Una lista de ENTITY -> ENTITIES

El nombre de una notación -> NOTATION

Existen dos atributos especiales: xml:lang y xml:space.

Tipo de atributo ID:

En una DTD, los atributos declarados ID son aquellos que solo pueden tomar un valor único (identificador) para cada elemento.

- Los valores de atributos ID, tienen que cumplir las mismas normas de sintaxis utilizadas para escribir nombres en XML. (no puede empezar con número)
- En un documento XML, no pueden escribirse dos elementos que tengan el mismo valor en un atributo ID, aunque dicho atributo sea distinto.
- Todo atributo declarado de tipo ID tiene que ser #IMPLIED (opcional) o #REQUIRED (obligatorio).