

XML DECLARACIÓN DEL TIPO DE DOCUMENTO DTD

DTD

- Establece las reglas de validación que debe cumplir un documento. (Recordatorio: diferencia entre documento bien formado y válido)
- Presenta deficiencias evidentes con respecto a XML Schema:
 - No tenemos un gran nivel de control sobre qué se considera válido.
 - No disponemos de control de tipos de datos (enteros, etc.).
 - No está definido como XML.

EJEMPLO DTD

PRÓLOGO DTD

- El prólogo se compone de la Declaración XML y la Declaración de Tipo de Documento.
 - Declaración XML
 - `<?xml version="1.0" encoding="ISO-8859-1" ?>`
 - Declaración de Tipo de Documento
 - Declaración DTD propiamente dicha
`<!DOCTYPE Casas_Rurales`
 - Definición de Tipo de Documento
 - Puede estar incluida dentro del documento.
`<!DOCTYPE Casas_Rurales [
...`
 - Puede ser externa al documento
 - `<!DOCTYPE NombreXML SYSTEM "URI" >`
 - `<!DOCTYPE NombreXML PUBLIC "id_publico" "URI" >`
 - Declaración de documento autónomo
 - `<?xml version="1.0" encoding="ISO-8859-1" standalone="no"?>`

DEFINICIÓN DE TIPO DE DOCUMENTO

- Declaraciones de tipos de elementos.
- Declaraciones de listas de atributos para los tipos de elementos.
- Declaraciones de entidades.
- Declaraciones de notación.

DECLARACIONES DE TIPOS DE ELEMENTOS

- **ESPECIFICACIÓN GENERAL**

- `<!ELEMENT nombre_elemento contenido>`
- `<!ELEMENT Dirección (#PCDATA) >`

- **ESPECIFICACIÓN DE CONTENIDO**

- **El tipo ANY:**

- Este tipo no impone ninguna restricción. (similar a AnyType de XSD)
 - `<!ELEMENT nombre_de_elemento ANY>`

- **El tipo EMPTY:**

- Elementos vacíos (similar a los elementos vacíos de XSD)
 - `<!ELEMENT nombre_de_elemento EMPTY>`

- **El tipo PCDATA:**

- Los elementos de dicho tipo contendrán únicamente datos carácter.
 - `<!ELEMENT nombre_de_elemento #PCDATA>`

- **El tipo CHILDREN:**

- Los elementos tienen otros elementos (similar a los elementos con otros elementos de XSD)
 - `<!ELEMENT nombre_elemento (elemento1, ... , elementoN)>`
 - `<!ELEMENT nombre_elemento (elemento1 | ... | elementoN)>`
 - `<!ELEMENT nombre_elemento (elemento1 | (elemento2,elemento3))>`

- **El tipo MIXED**

- El contenido mixto permite mezclar datos carácter y subelementos (similar a los elementos mixed de XSD)
 - `<!ELEMENT nombre_de_elemento (#PCDATA|Elemento1|Elemento2|...)*>`

TIPOS DE ELEMENTOS CHILDREN I

- **Secuencia de elementos**

- Se especifica la lista de los elementos que puede contener separados por comas.
 - *<!ELEMENT nombre_elemento (elemento1, ... , elementoN)>*

- **Alternativa de elementos |**

- Se utiliza como separador la barra vertical | .
- Únicamente un elemento de la lista puede formar parte del contenido en cada aparición del tipo de elemento.
 - *<!ELEMENT nombre_elemento (elemento1 | ... | elementoN)>*

- **Combinación de modelos**

- Se pueden utilizar paréntesis para agrupar secuencias y alternativas, y así conseguir combinaciones de los modelos anteriores.
 - *<!ELEMENT nombre_elemento (elemento1 | (elemento2,elemento3))>*
 - *<!ELEMENT nombre_elemento (elemento1 , (elemento2| elemento3))>*

TIPOS DE ELEMENTOS

CHILDREN II

- Especificación de frecuencia

?	opcional (0 o 1 vez)	Indica opción, el elemento o grupo se puede repetir cero o una vez
*	opcional y repetible (0 o más veces)	Indica cero o más repeticiones
+	necesario y repetible (1 o más veces)	Indica una o más repeticiones. Asegura una ocurrencia como mínimo

No es posible especificar una frecuencia determinada, como 3, 4, 7 veces.

Si no se especifica ninguno de estos caracteres el significado es de 1 y como máximo 1.

EJERCICIOS ELEMENTOS DTD

ATRIBUTOS

- Los atributos no pueden contener subatributos (los elementos pueden contener subelementos, lo que permite estructuración de la información)
- Los atributos solo se pueden especificar una vez y en cualquier orden.
- Tipos de Atributos
 - **Tipos cadena (CDATA)**
 - `<!ATTLIST nom_elemento nom_atributo CDATA valor_por_defecto>`
 - `<!ATTLIST perro fecha_nacimiento CDATA "">`
 - **Tipos enumerados**
 - `<!ATTLIST elemento atributo(valor1 | valor2 |...) valor_defecto>`
 - `<!ATTLIST mensaje prioridad (normal | urgente) normal>`
 - **Tipos notation**
 - Este atributo indica que el contenido del elemento debe ser procesado mediante la herramienta que indica la notación. Se verá más adelante.
 - **Tipos específicos (tipo token)**
 - NMTOKEN y NMTOKENS
 - ID, IDREF, IDREFS
 - ENTITY, ENTITIES

DECLARACIÓN POR DEFECTO

- **#REQUIRED** El atributo se ha de especificar obligatoriamente en el elemento.
- **#IMPLIED** El atributo no se ha de especificar obligatoriamente en el elemento.
- **#FIXED** "valor del atributo" El atributo solo puede contener el valor especificado
- **Un atributo puede incluir en su declaración valores por defecto.**
- Ejemplos
 - `<!ATTLIST perro fecha_nacimiento CDATA #REQUIRED>`
 - `<!ATTLIST perro madre IDREF #IMPLIED>`
 - `<!ATTLIST perro padre IDREF #IMPLIED>`
 - `<!ATTLIST perro otrosnombres NMTOKENS "">`
 - `<!ATTLIST perro veterinario CDATA #FIXED "Felix Marquez Sanz">`

EJERCICIOS DE ATRIBUTOS

ATRIBUTOS TIPO TOKEN

- Los atributos de tipo CDATA que puede contener casi cualquier carácter.
- Los atributos de tipo token debe contener un nombre XML y sólo puede contener:
 - letras dígitos . (punto) – (guión) _ (subrayado) : (dos puntos) y cumplir con la regla de producción número [7] de la Recomendación de XML.
- **NMTOKEN**
 - Sus valores válidos solo pueden ser Token.
 - <!ATTLIST mensaje fecha NMTOKEN #REQUIRED>
 - <mensaje fecha="15-7-2012">
- **NMTOKENS**
 - Los atributos del tipo NMTOKENS es una lista de NMTOKEN separados con espacios en blanco. Es decir aceptan varios valores NMTOKEN en un mismo atributo.
 - Un espacio en blanco consiste en uno o más espacios, retornos de carro o tabuladores
 - <!ATTLIST mensaje fecha NMTOKENS #REQUIRED>
 - <mensaje fecha="quince de Julio del dosmildoce">.

ATRIBUTOS TIPO TOKEN II

- **ID**

- Sirve para identificar al elemento dentro del documento de manera única. Un elemento solo puede tener como máximo un atributo de tipo ID.

- **IDREF**

- Se utilizan para referenciar elementos identificados con un atributo ID.
- Por tanto el valor de los atributos IDREF contiene el valor de otro atributo ID.

- **IDREFS**

- Los atributos IDREFS son análogos a los IDREF con la diferencia de que pueden tomar varios valores, separados por espacios.

- **Ejemplo**

- `<!ELEMENT enlace EMPTY>`
- `<!ATTLIST enlace destino IDREF #REQUIRED>`
- `<!ELEMENT capitulo (parrafo)*>`
- `<!ATTLIST capitulo referencia ID #IMPLIED>`

EJERCICIOS DE ATRIBUTOS 2

ENTIDADES Y NOTACIONES

- **Entidades generales internas**

- Son básicamente abreviaturas definidas en la sección de la DTD del documento XML.

```
<!DOCTYPE texto[  
  <!ENTITY fpe "Formación Profesional Específica">  
>
```

- XML tiene 5 entidades internas predefinidas: <>&'"

- **Entidades generales externas**

- Obtienen su contenido en cualquier otro sitio del sistema, ya sea otro archivo del disco duro, una página web o un objeto de una base de datos.
- Se hace referencia al contenido de una entidad así mediante la palabra SYSTEM seguida de un URI (Universal Resource Identifier)

ENTIDADES GENERALES EXTERNAS

- **Entidades generales externas no analizadas**

- Si el contenido de la entidad es un archivo MPG o una imagen GIF o un fichero ejecutable EXE, el procesador XML no debería intentar interpretarlo como si fuera texto XML.
- Las notaciones se usan para indicar el tipo de los ficheros binarios y cuál es el programa que debe interpretarlo.
- Ejemplo:

```
<!NOTATION jpg SYSTEM "programadegraficos.exe">
<!NOTATION gif SYSTEM "programadegraficos.exe">
<!ENTITY P250-01 SYSTEM "fotos/250-01.JPG" NDATA
jpg>
<!ENTITY P250-02 SYSTEM "fotos/P250-02.GIF"
NDATA gif>
<!ENTITY P250-03 SYSTEM "fotos/P250-03.JPG"
NDATA jpg>
<!ATTLIST perro fotos ENTITIES #IMPLIED>
...
```

- **Entidades generales externas analizadas**

```
<!DOCTYPE libro system "libro.dtd"
[
<!ENTITY indice SYSTEM "indice.xml">
<!ENTITY cap1 SYSTEM "capitulos/cap1.xml">
<!ENTITY cap2 SYSTEM "capitulos/cap2.xml">
]>
```

```
<libro>
  <cabecera>
    &indice;
  </cabecera>
  <cuerpo>
    &cap1;
    &cap2;
  </cuerpo>
</libro>
```

TIPOS DE ENTIDADES: ENTIDADES PARÁMETRO

- Entidades parámetro internas y externas
 - Se denominan entidades parámetro a aquellas que sólo pueden usarse en la DTD, y no en el documento XML.
 - Se pueden utilizar para agrupar ciertos elementos del DTD que se repitan mucho.
 - Se diferencian las entidades parámetro de las generales, en que para hacer referencia a ellas, se usa el símbolo % en lugar de & tanto como para declararlas como para usarlas.

- Ejemplo:

- Antes de saber manejar entidades con parámetros:

```
<!ELEMENT motos EMPTY>
```

```
<!ATTLIST motos a CDATA #REQUIRED b CDATA #REQUIRED c CDATA #REQUIRED>
```

- Después de saber manejar entidades con parámetros:

```
<!ENTITY %items "a CDATA #REQUIRED b CDATA #REQUIRED c CDATA #REQUIRED">
```

```
<!ELEMENT motos EMPTY>
```

```
<!ATTLIST motos %items>
```

REPASO ENTIDADES XSD