

XML

Qué significa el acrónimo?

Inglés → Extensible Markup Language

Español → Lenguaje de Marcado Extensible

Es un lenguaje que **almacena información** a partir de unas etiquetas **creadas por el usuario**

Este lenguaje hay que comprobar si está bien formado, para ello:

- DTD
- esquemas XML

¿Qué significan esas siglas?

Document Type Definition

Definición de tipo de documento

¿Qué es un DTD?

Es una declaración de la estructura y sintaxis de un documento XML.

Todos los documentos que usen la misma DTD mantendrán la consistencia.

DTD

Antes de crear un documento XML deberemos analizar que elementos y etiquetas se van a necesitar para la aplicación que se quiere crear.

- DTD se puede ubicar dentro del propio documento XML
- DTD se puede ubicar fuera del documento XML (fichero externo)

Ambas realizan la misma función, es cuestión de comodidad.

BLOQUES PARA CONSTRUIR DTD

- Elemento. Bloque principal.
- Atributo. Añadimos información al elemento
- Entidad. Caracteres especiales.
- PCDATA. Parsed Character Data. Indica que entre la apertura y cierre de este elemento se almacenarán caracteres como texto y serán analizados por un parser.
- CDATA. Character Data. **Caracteres que no necesitan ser analizados por el procesador.**

ELEMENTOS

- Se declaran de la siguiente manera:

<!ELEMENT nombre_del_elemento categoria>

O

<!ELEMENT nombre_del_elemento (nodos hijos)>

En este segundo caso, el nombre_del_elemento puede definirse como el nodo padre del que cuelgan un conjunto de nodos_hijos, separados por comas.

ELEMENTOS

- Para definir elementos vacíos, existe la palabra reservada EMPTY, que si se sitúa en la zona de categoría permite indicar que el elemento está vacío.

`<!ELEMENT saltoLinea EMPTY>` → `<saltoLinea />`

Dentro de los elementos se puede indicar si no están compuestos por nodos hijos, si son vacíos (EMPTY) o si contienen información (PCDATA, CDATA, ANY).

`<!ELEMENT teléfono (#PCDATA)>`

`<!ELEMENT hora (#CDATA)>`

Si especificáramos #ANY cualquier elemento conocido sería válido.

CANTIDADES Y ELECCIONES DE ELEMENTOS

CANTIDADES

- `<!ELEMENT nombre_del_elemento (categoria)>` = 1 y solo 1
- `<!ELEMENT nombre_del_elemento (categoria?)>` = Ninguna o 1
- `<!ELEMENT nombre_del_elemento (categoria+)>` = 1 o más veces
- `<!ELEMENT nombre_del_elemento (categoria*)>` = Ninguna o muchas

ELECCIONES

- `<!ELEMENT nombre_del_elemento (categoria A | categoria B)>`

DTD EJEMPLO

Ejemplo que utiliza DTD. Se quiere almacenar los mensajes de móviles que se envían a un servidor. Se guardarán estos datos:

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

```
<!ELEMENT sms (telefono, fecha, hora, mensaje)> Definimos elemento padre sms.  
<!ELEMENT telefono (#PCDATA)> Definimos los elementos hijos de sms.  
<!ELEMENT fecha (#PCDATA)>  
<!ELEMENT hora (#PCDATA)>  
<!ELEMENT mensaje ANY> ANY y EMPTY van sin paréntesis.  
<!ELEMENT saltolinea EMPTY>
```

ESTO ES EL DTD

Una vez definido el léxico, faltaría albergar las etiquetas con sus elementos, que pertenecería al XML, el siguiente paso.

DTD ejemplo 2

El elemento SMS esta compuesto por TELEFONO, FECHA, HORA y MENSAJE. Es una relación de padre-hijos.

Imaginemos que por cada SMS se pueden almacenar varios mensajes:

```
<!ELEMENT sms (telefono, fecha, hora, mensaje+)>  
<!ELEMENT telefono (#PCDATA)>  
<!ELEMENT fecha (#PCDATA)>  
<!ELEMENT hora (#PCDATA)>  
<!ELEMENT mensaje ANY o (#PCDATA)>  
<!ELEMENT saltolinea EMPTY>
```

El más para indicar 1 o más veces.

Porque són de caracteres texto.

ATRIBUTOS

- Se definen de la siguiente manera (sintaxis):

`<!ATTLIST nombre_elemento nombre_atributo categoría_atributo valor_por_defecto>`

- Nombre_elemento: Es el elemento al que se le quiere añadir el atributo.
- Nombre_atributo: es el nombre del atributo que se quiere añadir.

`<!ATTLIST hora zona CDATA "GMT+1">`

ATRIBUTOS

• **Tipo atributo:** existen muchos tipos, principales:

• **CDATA:** Texto, podrán tener cualquier carácter.

• **ID:** es un identificador que permite identificar al elemento de manera única.

• **IDREF:** es un identificador de otro elemento del propio documento XML.

• **IDREFS:** es una lista de identificadores a otros elementos (tipo 1 | tipo 2 | tipos | ...)

• **NMTOKEN:** es un texto que solo podrá tener letras, dígitos, guión, subrayado, punto y dos puntos. Lo que se define como nombres válidos XML.

• **NMTOKENS:** es una lista de nombre XML válidos. Es como un NMTOKEN pero se incluyen los espacios en blanco, tabuladores o retornos de carro.

• **ENTITY:** el tipo del atributo es una entidad que se ha declarado anteriormente

• **ENTITIES:** Es una lista de entidades.

ATRIBUTOS

- Los atributos podrían ser declarados como obligatorios **#REQUIRED**, optativos **#IMPLIED** o fijos **#FIXED**

En el DTD aplicaríamos:

```
<!ATTLIST hora zona CDATA "GMT+1" #REQUIRED>
```

Ejemplo

Construye un DTD que permita almacenar una base de datos (bbdd) de empleados, con las siguientes características.

La bbdd de empleados tendrá los campos: El departamento de RRHH que realiza los perfiles, solicitudes de perfiles (al menos uno), teléfono fijo o móvil, fecha y opcionalmente decir la ubicación. El departamento RRHH tendrá un atributo que será un código como identificador único, además de otros campos como email, teléfono, trabajadores (al menos uno)

El elemento Ubicación se compone de los elementos calle y portal. También tendrá dos atributos: numUbicación que será el identificador único y la ciudad que será opcional.

Falta terminarlo

Ejemplo

C:\Users\eperez\Desktop\Edgar Pérez DAM-DAW\Lenguaje de marcas y sistemas (MP4)\a.dtd • - Sublime Text (UNREGISTERED)

File Edit Selection Find View Goto Tools Project Preferences Help

```
a.dtd
1 <!ELEMENT sms (telefono, fecha, hora, mensaje+)>
2 <!ELEMENT telefono (#PCDATA)>
3 <!ELEMENT fecha (#PCDATA)>
4 <!ELEMENT hora (#CDATA)>
5 <!ELEMENT mensaje (#PCDATA)>
6 <!ELEMENT saltolinea EMPTY>
7
8
9
10
11 <!ELEMENT bbdd(empleados+)>
12 <!ELEMENT empleados (RRHH, fijo|movil, fecha, ubicacion?)>
13 <!ELEMENT RRHH (email,telefono_RRHH, trabajadores_RRHH+)>
14 <!ATTLIST RRHH codigo ID #REQUIRED>
15
16
```

DUDAS

