DAM

Lenguajes de Marcas

Tema 1: XML

XML (Extensive Markup Language) está desarrollado por W3C (World Wide Web Consortium).

Permite almacenar e intercambiar datos, siendo un predecesor de HTML. Tiene como formato el texto plano. Está compuesto por etiquetas.

<nombre>Valor</nombre>

También hay etiquetas sin valor o vacías (pueden llenarse más tarde), que se indican como <nombre/>. Ambas formas de escribir son equivalentes.

Puede haber elementos anidados.

<persona>

<nombre>Elsa</nombre>

<edad>25</edad>

</persona>

Se pueden mezclar etiquetas.

<nombre>Elsa</nombre> tiene <edad>25</edad> años.

Las etiquetas son case-sensitive, no es lo mismo escribir persona que Persona. Siempre hay que abrir y cerrar igual.

Los caracteres permitidos en XML son mayúsculas, minúsculas, números, -, \_. Se usan dos puntos para espacios de nombres.

El primer carácter debe ser una letra o un guión bajo. Las letras no inglesas, como la ñ o las acentuadas (con tilde) se pueden utilizar, pero no se recomienda por problemas de compatibilidad.

Algunos ejemplos de etiquetas mal escritas: <2color>, <persona><nombre> </persona></nombre>, <tag><tag/>.

**Atributos**

Dan información extra a las etiquetas. Se escriben tras la etiqueta inicial. Primero se le da un nombre al atributo y luego se escribe = y " ".

<producto color="azul"></producto>

No se pueden poner valores distintos a un mismo atributo. En el ejemplo anterior, si al lado ponemos color="rojo" no podemos ponerlo porque se sobreescribe. Hay que usar otro atributo o escribirlo de forma distinta.

<dato x="3" x="4"> está mal. En su lugar debemos poner <dato x="3" X="4"> o <dato x="3" y="4">.

XML es un lenguaje interpretado. Esto quiere decir que va línea a línea y que no hay programas que compilen.

Todo documento XML debe empezar con esta línea de código (declaración):

<?xml version="1.0" encoding="UTC-8"?>

Guardamos código xml con la extensión .xml y seleccionando *todos los archivos*.

En XML, por la codificación de los teclados, algunos caracteres se reconocen y otros no. Si queremos poner algo que no está en todos los idiomas hay que poner las llamadas **entidades de relación**.

&lt; menor que <

&gt; mayor que >

&quot; comilla doble “

&apos; comilla simple ‘

&amp; ampersand &

Si colocamos directamente los símbolos podrían no reconocerse, por eso se ponen estos mini códigos.

Comentarios: <!--Así escritos. No permiten -- dentro, ni anidaciones, ni escribir etiquetas.-->

**CData**

Almacenan grandes cantidades de texto. Puedes poner cosas en otro idioma, aislándolo del resto del código para que no haya problemas.

<![CDTA[Lorem Ipsum]]>

**Espacios de nombres**

A veces, una misma etiqueta puede ser ambigua. ¿Cartas de naipe o de menú? Podemos dar nombres distintos para diferenciar entre ambos.

<e1:carta xmlns:e1=”web\_inventada”

e2:carta xmlns:e2=”web\_inventada\_dos”>

Las urls no tienen que contener nada, solo sirven para guardar los espacios.

<e1:carta>As</e1:carta>

<e2:carta>Postres</e2:carta>

Para asignar valores por defecto, poner xlmns=”valor\_que\_se\_quiera”. Para ponerlos vacíos, simplemente poner comillas: xlmns= “”.

**Conservar espacios**

A veces nos puede convenir conservar tabulaciones y espacios, ya que XML los omite. Para conservarlos, hay que usar el atributo **xml:space=”preserve”**.