

1.XML

XML: eXtensible Markup language.

Surge en 1998 a partir del SGML y del HTML.

Su principal objetivo es almacenar la información y no mostrarla.

El lenguaje XML incluye una serie de archivos obligatorios, cada uno con una función. El resultado final consiste, de alguna manera, en la elaboración de una página HTML:

- El archivo XML que contiene el conjunto de datos (.xml)
- La DTD (Document Type Definition) contiene la definición de las características de los elementos que se van a mostrar (.dtd)
- Los archivos CSS (Cascading Style Sheet) permiten la presentación de los datos (.css)
- Los archivos XSL (XML Stylesheet Language) son hojas de estilo mucho más poderosas que las CSS y se adaptan mejor al XML (.xsl).

1.1. MI PRIMER DOCUMENTO XML

Un archivo XML es un archivo de texto con extensión .xml que contiene un determinado número de etiquetas estructuradas en árbol.

Un documento XML puede editarse con cualquier editor de texto y es interpretado por los navegadores Web.

Ejemplo, editamos con xmllcopy el siguiente documento y lo guardamos con extensión .xml:

```
<?xml version="1.0"?>
```

```
<tarjetavisita>
```

```
    <apellido>DEL VALLE</apellido>
```

```
    <nombre>Cristobal</nombre>
```

```
    <sociedad/>
```

```
    <profesion>Doctor</profesion>
```

```
    <direccion>
```

```
        <numero>19</numero>
```

<calle>C/ Florida</calle>

<codigopostal>28080</codigopostal>

<poblacion>MADRID</poblacion>

</direccion>

<numerotelefono>123456789</numerotelefono>

<numeromovil/>

<numerofax/>

<email/>

</tarjetavisita>

Si al abrir el documento el navegador nos dice:

No se puede mostrar la página XML...

Es porque contiene errores.

1.2. LA PRIMERA LINEA DE UN DOCUMENTO XML

Declaración XML

La primera línea de un documento XML se denomina prólogo y contiene, la declaración de la versión XML empleada para la descripción de la información. Tiene la siguiente forma:

```
<?xml versión="1.0" ?>
```

Codificación de los caracteres

En el prólogo podemos especificar el estándar de la codificación de caracteres utilizados en el documento. Con UTF-8 puedo representar los acentos y la ñ.

```
<?xml versión="1.0" encoding="UTF-8"?>
```

Validación y estilos

También contendrá los vínculos a los documentos DTD, CSS y XSL, lo veremos más adelante.

1.3. LOS ELEMENTOS

Constituyen la arquitectura en árbol de un documento XML. El nombre de los elementos se encuentra en el interior de las etiquetas y su valor se sitúa entre una etiqueta de apertura y una etiqueta de cierre:

`<NombreElemento>Valor del elemento</NombreElemento>`

Elementos vacíos

Un elemento vacío es un elemento cuyo valor no ha sido definido. Existen dos formas de especificar un elemento vacío:

`<elemento></elemento>`

O

`<elemento/>`

Mayúsculas y minúsculas

Los nombres de los elementos deben ser idénticos en la apertura y el cierre, distinguiéndose entre mayúsculas y minúsculas.

Caracteres permitidos en el nombre de un elemento

No se permiten los espacios.

Se pueden utilizar:

- Los siguientes signos: -, : (mejor evitarlo), _ , .
- Todos los caracteres alfabéticos acentuados o no.
- Los caracteres numéricos.

El nombre de un elemento no puede comenzar por número y además se debe evitar que el nombre comience por XML o por x (X) m (M) l (L).

Atributos

Son informaciones adicionales asociadas a los elementos.

`<elemento atributo="valor del atributo">Valor del elemento</elemento>`

Ejemplo:

`<persona sexo="masculino">DEL VALLE</persona>`

No se deben utilizar en exceso, siempre que se pueda se deberían convertir en elementos.

1.4. RAÍZ DE UN DOCUMENTO

Todos los documentos XML deben contener una raíz, que es el elemento padre del resto de elementos.

En el siguiente ejemplo la raíz es directorio:

```
<?xml version="1.0" encoding="UTF-8"?>
```

```
<directorio>
```

```
<entrada>
```

```
<apellido>BENITEZ</apellido>
```

```
<nombre>Javier</nombre>
```

```
<fechanacimiento/>
```

```
<dirección/>
```

```
</entrada>
```

```
<entrada>
```

```
<apellido>MENÉNDEZ</apellido>
```

```
<nombre>Miguel</nombre>
```

```
<fechanacimiento>18/2/1968</fechanacimiento>
```

```
<dirección/>
```

```
</entrada>
```

```
<entrada>
```

```
<apellido>LÓPEZ</apellido>
```

```
<nombre>Juan</nombre>
```

```
<fechanacimiento/>
```

```
<dirección/>
```

```
</entrada>
```

```
</directorio>
```

1.5.COMENTARIOS

Información que no es interpretada por el navegador:

<!-- Esto no lo interpreta el navegador -->

1.6. ESPACIO DE NOMBRES

Un **espacio de nombres XML** sirve para proporcionar elementos y atributos con nombre único en un documento XML. Así se resuelve la ambigüedad existente entre elementos o atributos que se llamen igual. Los nombres de elementos dentro de un espacio de nombres deben ser únicos.

Un ejemplo sería un documento XML que contuviera referencias a un cliente y a un producto solicitado por éste. Tanto el elemento que representa el cliente como el que representa el producto pueden tener un elemento hijo llamado "numero_ID". Las referencias al elemento "numero_ID" podrían ser ambiguas, salvo que los elementos, con igual nombre pero significado distintos, se llevaran a espacios de nombres distintos que los diferenciaran.

```
<?xml version="1.0"?>
<cli:cliente xmlns:cli="http://www.mipagina/Espacio_de_nombres_XML/cliente"
  xmlns:ped="http://www.mipagina/Espacio_de_nombres_XML/pedido">
  <cli:numero_ID>1232654</cli:numero_ID>
  <cli:nombre>Fulanito de Tal</cli:nombre>
  <cli:telefono>99999999</cli:telefono>
  <ped:pedido>
    <ped:numero_ID>6523213</ped:numero_ID>
    <ped:articulo>Caja de herramientas</ped:articulo>
    <ped:precio>187,90</ped:precio>
  </ped:pedido>
</cli:cliente>
```

Dentro de un elemento es posible hacer referencia a un espacio de nombres utilizando la siguiente sintaxis:

```
<nombre_elemento xmlns:prefijo="URI">
```

Dónde URI (Universal Resource Identifier) es la dirección del documento del espacio de nombres.

Dentro de un documento para utilizar elementos con el mismo nombre, se hace preceder el nombre del elemento o del atributo por el prefijo del espacio de nombres seguido de : (dos

puntos).

Sin embargo, hay que destacar que **el URI no se lee realmente como una dirección**; se trata como una cadena de texto. **El hecho de usar una URI para identificar un espacio de nombres**, en lugar de una simple cadena (como "xhtml"), **reduce la posibilidad de que diferentes espacios de nombres usen identificadores iguales**. Los identificadores de los espacios de nombres no necesitan seguir las convenciones de las direcciones de internet, aunque a menudo lo hagan.

Un espacio de nombres XML no necesita que su vocabulario sea definido, aunque es una buena práctica utilizar un DTD o un esquema XML para definir la estructura de datos en la ubicación URI del espacio de nombres. (Lo veremos más adelante).

El alcance de la declaración de un prefijo de espacio de nombres comprende desde la etiqueta de inicio de un elemento XML, en la que se declara, hasta la etiqueta final de dicho elemento XML. En las etiquetas vacías, correspondientes a elementos sin "hijos", el alcance es la propia etiqueta.

1.6. DOCUMENTO XML BIEN FORMADO

Un documento XML no está "bien formado", si no respeta las siguientes normas:

- El prólogo no es válido.
- Algunas etiquetas no se han cerrado.
- Los elementos no están conectados correctamente.
- Los valores de los atributos no se han colocado entre comillas.
- Los nombres de los elementos contienen caracteres no permitidos.
- No existe ningún elemento raíz en el documento.