



## Índice

# Conceptos básicos de XML

¿Qué es lenguaje de marcas?

Llamamos lenguaje de marcas a la forma de codificar un documento incorporando etiquetas o marcas que contienen información adicional acerca de la estructura del texto o de su presentación. El lenguaje de marcas más extendido y que es fundamento de la Web es el HTML.

Dos diferencias fundamentales entre HTML y XML?

Se diferencian en su **objetivo**: XML sirve para describir información y el HTML sirve para darle formato y presentarla a través de un navegador.

Y también en su **sintaxis**: HTML se utilizan etiquetas estándar, mientras que en XML el usuario puede crear sus propias etiquetas de descripción.

¿Qué es XML?

XML son las siglas de eXtensible Markup Language (Lenguaje de Marcas Extensible). Es un "lenguaje" de etiquetas desarrollada por el **W3C**, diseñado especialmente para los documentos de la web. Consigue que los diseñadores creen sus propias etiquetas, permitiendo la definición, transmisión, validación e interpretación de datos entre aplicaciones y entre organizaciones.

¿En qué formato se escriben los ficheros XML?

Los ficheros XML se escriben en formato texto plano, con una sintaxis específica para organizar la información en estructura de árbol.

Texto plano es aquel en el que la única información es la de los caracteres, sin formato ni ningún otro añadido.

¿Qué es el prólogo de un documento XML? ¿Qué elementos contiene?

Aunque no es obligatorio, los documentos XML pueden empezar con unas líneas que describen la versión XML, el tipo de documento y otras informaciones referidas al documento, denominada Prólogo.

Además de la declaración XML, puede contener la codificación de caracteres, la estructura del documento, las hojas de estilos, instrucciones de procesamiento y comentarios. Ejemplo:

```
<?xml version="1.0" encoding="utf-8"?>
<?xml-stylesheet type="text/xsl" href="show_book.xsl"?>
<!DOCTYPE catalog SYSTEM "catalog.dtd">
<!--catalog última actualización 2010-11-01-->
```

¿Cómo se introducen comentarios en un documento XML?

Igual que en HTML, el texto que queramos poner como comentario se coloca entre las marcas `<!--` y `-->`

Como ejemplo puedes ver la última línea de la pregunta anterior.

a) ¿Qué es un elemento? b) ¿Cómo se define?

- a) Un elemento es el componente básico o unidad de información de un documento XML.
- b) Se define con un nombre, que se encuentra en el interior de las etiquetas y su valor se sitúa entre una etiqueta de apertura y una etiqueta de cierre.

Ejemplo: `<nombre-elemento> Valor </nombre-elemento>`

a) ¿Qué es un elemento vacío? b) ¿Cómo se define?

- a) Un elemento vacío es un elemento cuyo valor no ha sido definido.
- b) Se puede definir de dos formas:

`<elemento></elemento>` ó simplemente `<elemento/>`

a) ¿Qué es la raíz? b) ¿Cómo se define? c) ¿Cuántas puede haber?

- a) Es el elemento del que parte el árbol que constituye el documento. Todos los demás dependen de la raíz. Sería el elemento inicial de una estructura en árbol.
- b) Se define como cualquier otro elemento.
- c) Sólo puede haber uno por cada estructura XML. Ejemplo:

```
<empresa>
  <empleado>
    <id>154</id>
    <nombre>Alejandro</nombre>
    <apellido>Magno</apellido>
    <dni/>
  </empleado>
</empresa>
```

En este ejemplo, la raíz sería el elemento `empresa`

¿Cómo se organizan los elementos de un documento XML?

Los documentos han de seguir una estructura (árbol) estrictamente jerárquica en lo que respecta a las etiquetas que delimitan sus elementos. Las etiquetas deben estar correctamente anidadas. Los elementos con contenido deben estar correctamente cerrados.

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

<raiz>
  <elemento1>
    <elemento11 atributo="contenido del atributo">
      <elemento111>contenido del elemento 111</elemento111>
      <elemento112>contenido del elemento 112</elemento112>
    </elemento11>
  </elemento1>
  <elemento2>Contenido del elemento2</elemento2>
</raiz>
```

¿Qué caracteres puede contener el nombre de un elemento?

Están permitidos todos los caracteres alfabéticos —acentuados o no— los caracteres numéricos, y el uso de los caracteres siguientes:

- (guion), \_ (guion bajo) y . (punto)

¿Cómo tiene que comenzar el nombre de un elemento?

El primer carácter del nombre de un elemento sólo puede ser un carácter alfabético o un guion bajo, y no debe comenzar por la cadena “xml” ni ninguna variante de la misma: “XML”, “Xml”, etc.

a) ¿Qué es un atributo? .....b) ¿Cómo se define?

c) Pon un ejemplo.

- a) Atributo es una información adicional asociada a un elemento.
- b) Se define dentro de la etiqueta de apertura, a continuación del nombre del elemento, mediante el par nombreDelAtributo="valor"
- c) Un ejemplo de definición de atributo sería:

```
<alumno nota="5" asistencia="Nula">Richalson</alumno>
```

a) ¿Qué es un espacio de nombres en XML? b) ¿Por qué conviene Utilizar Espacios de nombres?

- a) Un espacio de nombres XML es una recomendación [W3C](#) para proporcionar elementos y atributos con nombre único en una instancia [XML](#). Es un conjunto de elementos y atributos definidos de forma única que pertenecen al mismo vocabulario.
- b) Es necesario porque un documento XML puede contener nombres de elementos o atributos procedentes de más de un vocabulario XML. Si a cada uno de estos vocabularios se le da un [espacio de nombres](#), 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.

¿Cómo se declara un espacio de nombres en nuestro documento XML?

Un espacio de nombres se declara usando el atributo reservado **xmlns**, cuyo valor debe ser un [identificador uniforme de recurso](#) (URI).

Por ejemplo: **xmlns="http://www.w3.org/2001/XMLSchema"**

Sin embargo, hay que destacar que el URI no se lee realmente como una dirección de internet; el [Parser](#) XML lo trata como una cadena de texto. Por ejemplo, el propio <http://www.w3.org/1999/xhtml> no contiene código alguno, simplemente describe el espacio de nombres [XHTML](#) a lectores humanos. El hecho de usar una URL como esta 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

¿Qué es un Alias de un espacio de nombres?

Un nombre corto que se le da al identificador del espacio de nombres, para que nos podamos referir a éstos de una forma más sencilla.

Por ejemplo: le vamos a dar al identificador anterior el alias de xshe.

**xmlns:xshe="http://www.w3.org/2001/XMLSchema"**

Ejemplo de uso de espacios de nombres?

En el ejemplo siguiente, estamos distinguiendo el **numero\_ID** correspondiente al cliente del mismo nombre correspondiente al pedido.

```
<?xml version="1.0"?>
<cliente xmlns:cli='http://Espacio_de_nombres_XML/cliente'
          xmlns:ped='http://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>
```