# XML

# 

Tipos de lenguaje:

1. JSON (JavaScrip Object Notation). Es un formato ligero para el intercambio de datos. Es una alternativa a XML en AJAX.
2. XML (eXtensible Markup Language). Es un **metalenguaje** extensible, que puede ser usado para definir otros lenguajes. Es el otro tipo de SGML (de el surgía HTML y XML, y de la unión de ambos surge XHTML). Fue desarrollado por el W3C.

Recordatorios:

* Son lenguajes descriptivos, también llamados semánticos, en el que las marcas indican qué es la información (detallan que es lo que se esta representando), pero no especifican cómo representarla.

Son flexibles y reutilizables.

* Las etiquetas están compuestas por el símbolo <, un texto y el símbolo >. Hay etiquetas de inicio y de fin. También se conocen por el nombre tags o marcas.

Ejemplo: **<nombre>**Maite**</nombre>**

* Los elementos son un componente o estructura mediante la cual se organiza el contenido o se indica que acciones se desencadenan. Están formadas por una **etiqueta de inicio, etiqueta de fin y lo que se encuentre en ambas etiquetas.** Pueden ser simples (que no tienen ningún elemento hijo), o compuestos, que son los que están formados por mas elementos (elementos hijos). También puede haber elementos vacíos, es decir, que no tienen contenido. La etiqueta de fin es puede o no aparecer en algunos casos.
* Los atributos, indican las propiedades del elemento. Se encuentra siempre dentro de la etiqueta de inicio, su valor suele ir entre comillas, y su valor es obligatorio. Ejemplo: <alumno **id=”25”**>Pepe</alumno>

## XML

XML es la abreviatura de eXtensible Markup Language, que significa lo siguiente:

Lenguaje: Conjunto de reglas que definen una sintaxis y una grmática.

Marcado: Método para escribir o incrustar metadatos

¿Qué es un metadato?

Es la información que podemos aportar sobre un conjunto de datos.

Puede haber diferentes interpretaciones de un mismo texto, por lo tanto, las etiquetas tendrán diferente nombre, según esa interpretación. Por ejemplo:

El siguiente texto puede tener diferentes interpretaciones

Juan Antonio

Bochinche

Cuatro Esquinas

30190

Versión 1:

<alumno>

<nombre>Juan Antonio</nombre>

<APELLIDO>Bochinche</APELLIDO>

<Dirección>Cuatro Esquinas</Dirección>

<Código\_Postal>30190</Código\_Postal>

</alumno>

Versión 2:

<mision\_007>

<te-llamaras>Juan Antonio</te-llamaras>

<contra.seña>Bochinche</contra.seña>

<objetivo>Cuatro Esquinas</objetivo>

<tiempo\_estimado>30190</tiempo\_estimado>

</mision\_007>

En estos ejemplos, ya podemos ver la estructura que tiene este lenguaje:

1. Todo cuelga de un mismo elemento raíz.
2. Las etiquetas de cierre y apertura tienen que ser idénticas
3. Las etiquetas no pueden empezar por números.
4. Bien anidado: Si se abre una etiqueta y esta contiene otra, la que primero se cierra, es la segunda que se abrió.
5. Hay una serie de caracteres admitidos.
6. Tiene estructura de árbol.

Mas ejemplos:

Etiquetas:

<noticia>

<lugar>  Alcobendas </lugar>

<fecha> 01/02/2022 </fecha>

<evento>

Comienza el tercer trimestre en el

módulo de Lenguajes de Marcas

</evento>

</noticia>

ELEMENTOS (etiqueta inicio + contenido + etiqueta fin)

<noticia>

<lugar>Alcobendas</lugar>

<fecha>01/02/2022</fecha>

<evento>

Comienza el tercer cuatrimestre en el

módulo de Lenguajes de Marcas

</evento>

</noticia>

Atributos:

<noticia>

<lugar>Alcobendas</lugar>

<evento **fecha**=”01/02/2022” >

Comienza el tercer cuatrimestre en el

módulo de Lenguajes de Marcas

</evento>

</noticia>

No hay espacio en el símbolo del < o > y la palabra de la etiqueta.

Restricciones

Los nombres de todos los elementos son case sensitive (sensible a mayúsculas y minúsculas), teniendo que cumplir las siguientes normas:

* El primer carácter tiene que ser una letra o un guión bajo “\_”
* Sólo pueden contener letras (con o sin tilde), números, puntos “.”, guiones medios “-”, guiones bajos y dos puntos.
* Los dos puntos están reservados para definir espacios de nombres
* No pueden contener ningún carácter espaciado.

Restricciones de los atributos.

* Los identificadores de los atributos se forman de la misma manera que los identificadores de los elementos.
* El valor de los atributos siempre irá entre comillas: Las comillas pueden ser simples o dobles, pero nunca se podrán mezclar. (En HTML, las comillas son opcionales).
* Nunca se repite el nombre de un atributo dentro del mismo elemento.
* Los atributos siempre se incluyen en la etiqueta de inicio. Ejemplo incorrecto:
* <alumno id=”532” id=”345”>

<nombre>Pablo</nombre>

<apellido>Pérez</apellido>

<fecha día=22 mes=’10” 1año=’1985’ />

</alumno> …

<!-- En la etiqueta de inicio de un elemento no haber dos atributos con el mismo nombre.

**Aplicar las reglas de definición de elementos a los nombres de los atributos** -->

* Un uso excesivo de atributos puede provocar que el documento XML sea menos legible, mas difícil de entender y, por lo tanto, difícilmente extensible. Para evitar esto siguen las siguientes reglas:
  + Los atributos **no** pueden contener otros elementos o atributos, como si sucede en los elementos.
  + Se suelen usar como “metadatos”. Por ejemplo, identificadores.
  + Nunca se puede repetir el nombre de un atributo, dentro del mismo elemento (examen):

<persona id=”rubia”  ~~id=”castaña”~~>

* + Las clases también son atributos, por lo que sólo puede contener una clase, y separados por comas, los nombres de las distintas clases que le afecten a esa etiqueta.