

1.1

Los lenguajes de marcas (también llamados lenguajes de marcado) son aquellos que combinan la información, generalmente textual, que contiene un documento con marcas o anotaciones relativas a la estructura del texto o a la forma de representarlo.

1.2

Como sabemos existen formatos binarios de archivos que han sido estándares de facto, para imagen JPEG, Documento PDF, video MPEG... Gracias a esto podemos compartir archivos entre distintos equipos.

Aunque se hayan creado formatos estándares sigue habiendo empresas que utilizan su propio formato binario ya que se adecua más a un software determinado. Por tanto surge un problema a la hora de compartir estos archivos propios.

Para buscar una solución se utilizó el texto ya que es un formato de archivo que cualquier dispositivo y sistema operativo es capaz de entender. Pero... ¿Como es esto posible? Todos sabemos que los archivos de texto sólo son capaces de almacenar texto plano sin indicar ningún formato, lo que no sabíamos, hasta el momento es que existe una manera de almacenar distintos datos dentro del documento. De ahí surge la idea de los lenguajes de marcas, esta idea consiste en que dentro del archivo donde se aloja el texto informativo coexistirá otro texto que se interpretará de manera especial, haciendo por ejemplo que muestre una imagen (``)

1.3

Los lenguajes de marcado se suelen confundir con lenguajes de programación. Sin embargo, no son lo mismo, ya que el lenguaje de marcado no tiene funciones aritméticas o variables, como poseen los lenguajes de programación. Los lenguajes de marcado son un modo de codificar un documento donde, junto con el texto, se añaden etiquetas con información adicional relativa a la estructura o formato del texto.

1.4

Los lenguajes de marcado se clasifican en tres grandes grupos:

1. Lenguajes orientados a presentación

El marcado de procedimientos está enfocado hacia la presentación del texto, sin embargo, también es visible para el usuario que edita el texto. El programa que representa el documento debe interpretar el código en el mismo orden en que aparece. Codifican cómo ha de presentarse el documento, por ejemplo, indicando que una determinada palabra debe presentarse en fuente itálica o que se debe dejar un espacio de 10 puntos al terminar el párrafo.

Algunos ejemplos de marcado de procedimientos son “nroff, troff, TeX....”

2. Lenguajes procedurales

En este tipo de lenguajes las etiquetas son también orientadas a presentación, pero se integran dentro de un marco procedural que permite definir macros (secuencias de acciones) y subrutinas. Entre los ejemplos más comunes de lenguajes procedurales podemos encontrar TeX, LaTeX y Postscript.

3. Lenguajes descriptivos

Este tipo de lenguajes no definen qué se debe hacer con un trozo o sección del documento, sino que por el contrario las marcas sirven para indicar qué es esa información, esto es, describen que es lo que se está representando. La mayoría de los lenguajes de marcas que se usan hoy en día se encuentran dentro de este grupo.

Un ejemplo de estos son SGML y XML

1.5

GML son las siglas de General Markup Language y fue el primer lenguaje que usaba marcas, creado por IBM en los años 60. No era un lenguaje de marca propiamente dicho, pero fue la semilla que los germinó. GML se usaba en el ámbito empresarial para la transmisión de información usando redes de computadores ya que, al tratarse solamente de texto, los documentos son independientes de la plataforma, sistema, programa...

SMLG es el descendiente directo de GML, sus siglas significan Standard Generalized Language. El objetivo de este lenguaje era también compartir información por parte de sistemas informáticos, pero tampoco consiguió asentarse del todo debido a que provocaba que el software que usaba SGML terminaba siendo excesivamente extenso y complejo.

1.6

HTML se popularizó rápidamente por aprender de los anteriores y solucionar los problemas que tenían los anteriores lenguajes. HTML es un lenguaje de marcas sencillo, con una estructura agradable y estructurada y del que no se requiere grandes conocimientos para utilizar. Admite archivos pequeños y tiene compatibilidades con todos los clientes (navegadores).

HTML es el lenguaje que se utiliza para crear las páginas web a las que se accede mediante internet". Más concretamente, HTML es el lenguaje con el que se "escriben" la mayoría de páginas web.

1.7

W3C en el 98 empezó con el desarrollo de XML (Extended Markup Language). Pretendían solucionar las carencias del HTML en lo que se respecta al tratamiento de la información.

Problemas del HTML como:

- El contenido se mezcla con los estilos que se le quieren aplicar.
- No permite compartir información con todos los dispositivos, como pueden ser ordenadores o teléfonos móviles.
- La presentación en pantalla depende del visor que se utilice.

Como vimos antes, XML es un lenguaje descriptivo que proporciona una plataforma para definir elementos para crear un formato y generar un lenguaje personalizado.

1.9

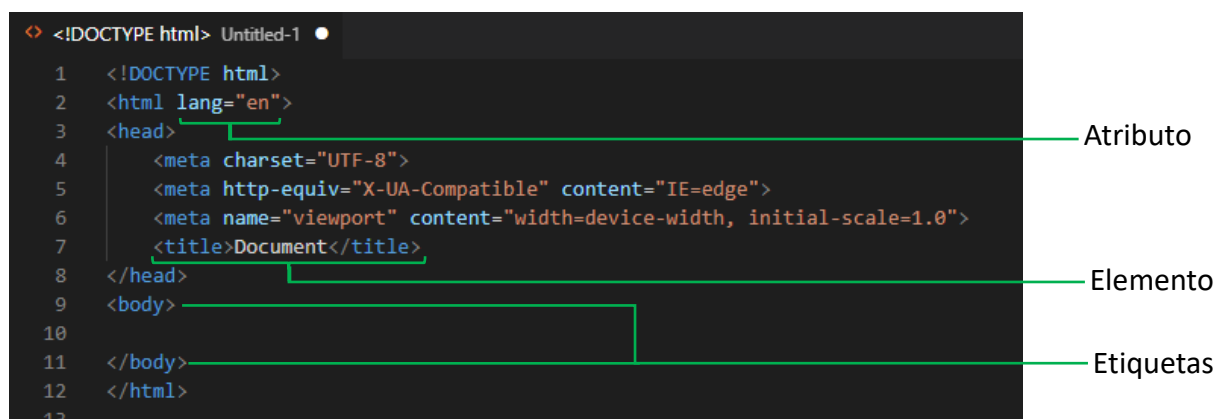
Según la [W3C](#):

-Las **etiquetas** son la estructura básica de HTML, siempre van entre signos de mayor que y menor que < >

-Los **atributos** van dentro de las etiquetas de inicio de un elemento y nos indica las propiedades que pueden llevar asociadas los elementos.

-Un **elemento** en HTML se refiere al conjunto de la etiqueta, su atributo (si tuviese), y el contenido que lo conforma.

Estos conceptos quedan más claros con una imagen de referencia:



1.10

Las características principales de los lenguajes de marcas son:

-Son un texto plano:

Una de las principales ventajas de los ficheros creados usando lenguaje de marcas es que pueden ser interpretados directamente, dado que son un archivo de texto plano y no necesitan de un programa intermediario para trabajar como pasa con los sistemas de archivos binarios. Esto también hace que sean independientes de la plataforma, SO o programa con el que fueron creados.

-Facilidad de procesamiento:

Los lenguajes de marcas son muy fáciles de procesar y esto permite que las organizaciones de estándares hayan desarrollado lenguajes especializados para diferentes tipos de documentos.

-Flexibilidad:

Aunque los lenguajes de marcas se crearon para documentos, a día de hoy se han empezado a utilizar en diferentes áreas como gráficos vectoriales, servicios Web o interfaces de usuario. Esto ha permitido que se puedan configurar varios lenguajes en una misma pagina para implementar audio, video, gráficos, formulas matemáticas...

1.11

La gramática es la ciencia que estudia los elementos de una lengua y sus combinaciones. Todo documento de un lenguaje de marcas tiene en común una gramática que define el marcado permitido, requerido y como debe usarse en ese documento en específico.

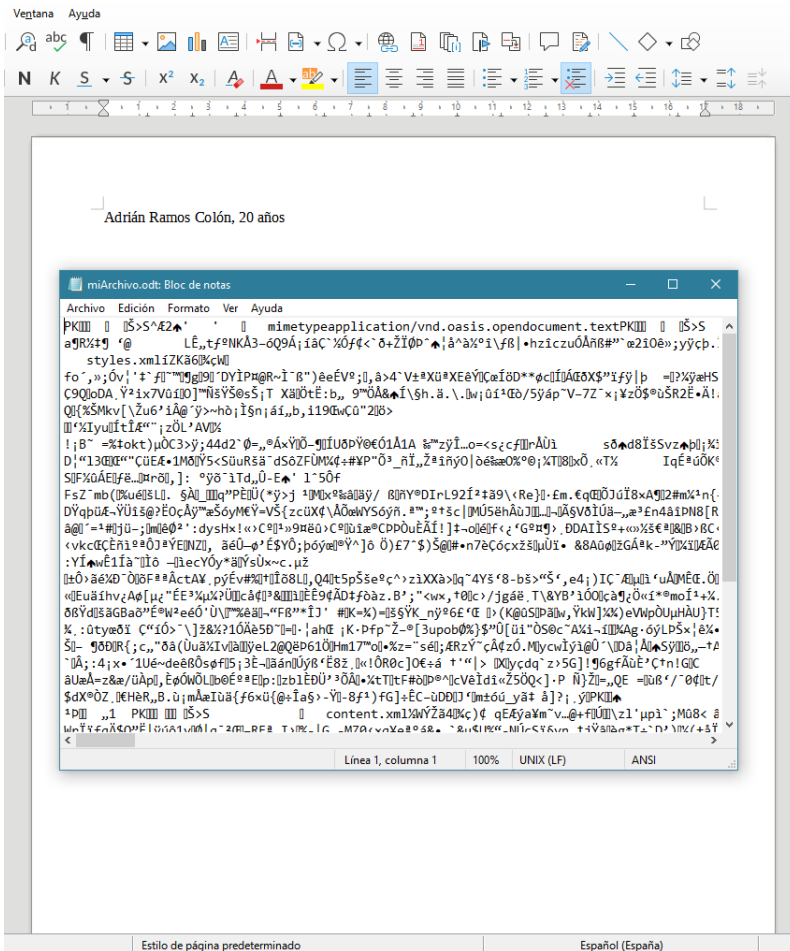
1.12

DTD: Establece las reglas de formación del lenguaje formal, es decir, qué combinaciones de símbolos elementales son sintácticamente correctas.

XML Schema: Es la evolución de la DTD descrita por el W3C. Es un lenguaje de esquema más complejo y más potente, basado en la gramática para proporcionar una potencia expresiva mayor que la DTD. Utiliza sintaxis XML que le permite especificar de forma más detallada un extenso sistema de tipos de datos.

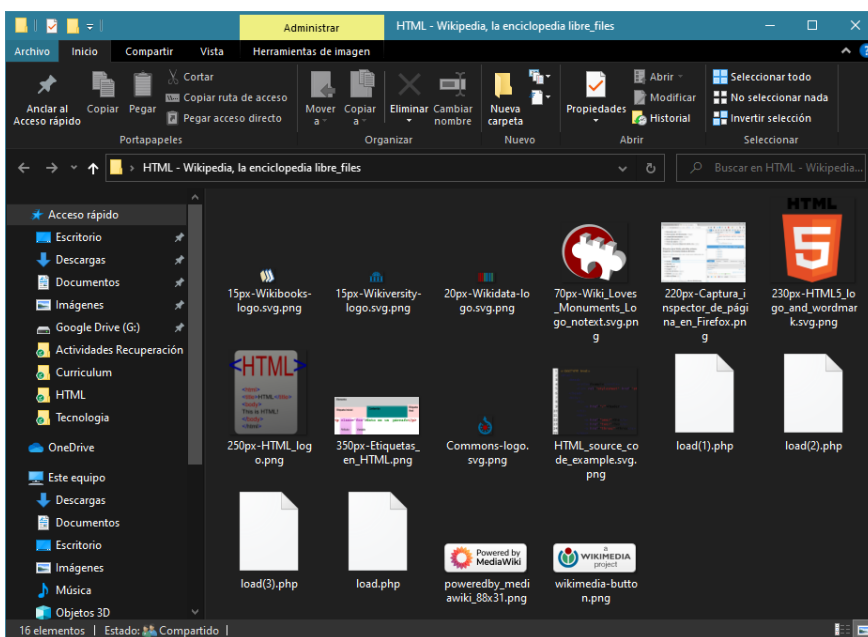
RELAX NG: Es un lenguaje de esquema basado en la gramática, muy intuitivo y más fácil de entender que el XML Schema. Tiene un alto poder expresivo, ya que, por ejemplo, permite validar elementos intercalados que pueden aparecer en cualquier orden.

2.



Como podemos observar en la imagen, al cambiar la extensión de .odt a .txt el archivo no es capaz de leer ya que no tienen la misma codificación y no son compatibles.

3.



Al guardar la página se crea una carpeta que contiene todos los archivos de la carpeta “public” de la web que te permite poder abrir la pagina en un navegador sin necesidad de internet.

4.

```
<?xml version="1.0" encoding="UTF-8"?>
<nitf>
  <head>
    <title>Colombia Earthquake</title>
  </head>
  <body>
    <headline>
      <h1>143 Dead in Colombia Earthquake</h1>
    </headline>
    <byline>
      <bytag>By Jared Kotler, Associated Press Writer</bytag>
    </byline>
    <dateline>
      <location>Bogota, Colombia</location>
      <date>Monday January 25 1999 7:28 ET</date>
    </dateline>
  </body>
</nitf>
```

Atributos

Elemento

Etiquetas

Este documento es un ejemplo de noticia donde se el titulo es Terremoto en Colombia, informa de las muertes producidas, el escritor del artículo, la localización y la fecha de la entrada. Como podemos observar, las etiquetas de XML no proporcionan ningún estilo, solo proporcionan información del contenido.

```
<?xml version="1.0" encoding="UTF-8"?>
<nitf>
  <head>
    <title>Colombia Earthquake</title>
  </head>
  <body>
    <headline>
      <h1>143 Dead in Colombia Earthquake</h1>
    </headline>
    <byline>
      <bytag>By Jared Kotler, Associated Press Writer</bytag>
    </byline>
    <dateline>
      <location>Bogota, Colombia</location>
      <date>Monday January 25 1999 7:28 ET</date>
    </dateline>
    <web title="Notice">
      <description
        xlink:type="simple"
        xlink:href="http://www.iespuntadelverde.es/ "
        xlink:show="new">
      </description>
    </web>
  </body>
</nitf>
```

En este caso hemos añadido una web donde podríamos añadir la url del noticiario web de la noticia.

5.

La pregunta es bastante ambigua, voy a mencionar las etiquetas y atributos más comunes:

```
<!DOCTYPE html>
<html></html>
<head></head>      <!--Estas etiquetas son las principales donde se sustenta
<title></title>      el resto de nuestro código-->
<body></body>

<blockquote></blockquote>
<footer></footer>
<div></div>          <!--Estas son las principales etiquetas para organizar
<title></title>      Y estructurar nuestra página -->
<nav></nav>
<header></header>

<p></p>
<h1></h1>            <!--Estas son las principales etiquetas para introducir
<h2></h2>...        texto y encabezados -->

<ul></ul>
<li></li>            <!--Estas son las principales etiquetas para introducir
<ol></ol>            listas ordenadas y desordenadas-->

    <!--Estos son los principales atributos junto con sus etiquetas más comunes-->

<img sizes="" src="" alt=""> sizes=tamaño, src=url externa, alt=texto alternativo
<a type="" href=""></a> type=tipo de dato, href=enlace de destino
<audio loop></audio> loop=bucle
<p id=""></p> id=identificación única
<h2 class=""></h2> class=especifica clase

    <!--Estos son los principales atributos que proporcionan información al navegador
    sobre nuestro documento-->

<meta charset=""> charset=codificación de caracteres
<meta http-equiv="" content=""> http-equiv=simula encabezado HTTP content=valor
                                asociado al atributo anterior
<meta name=""> name=especifica nombre
```