

LMSGI01.1 Lenguajes de marcas



☰ Intercambio de la información en la Web

☰ Introducción al lenguaje de marcas

☰ Etiquetas o marcas

☰ Evolución histórica

☰ Características

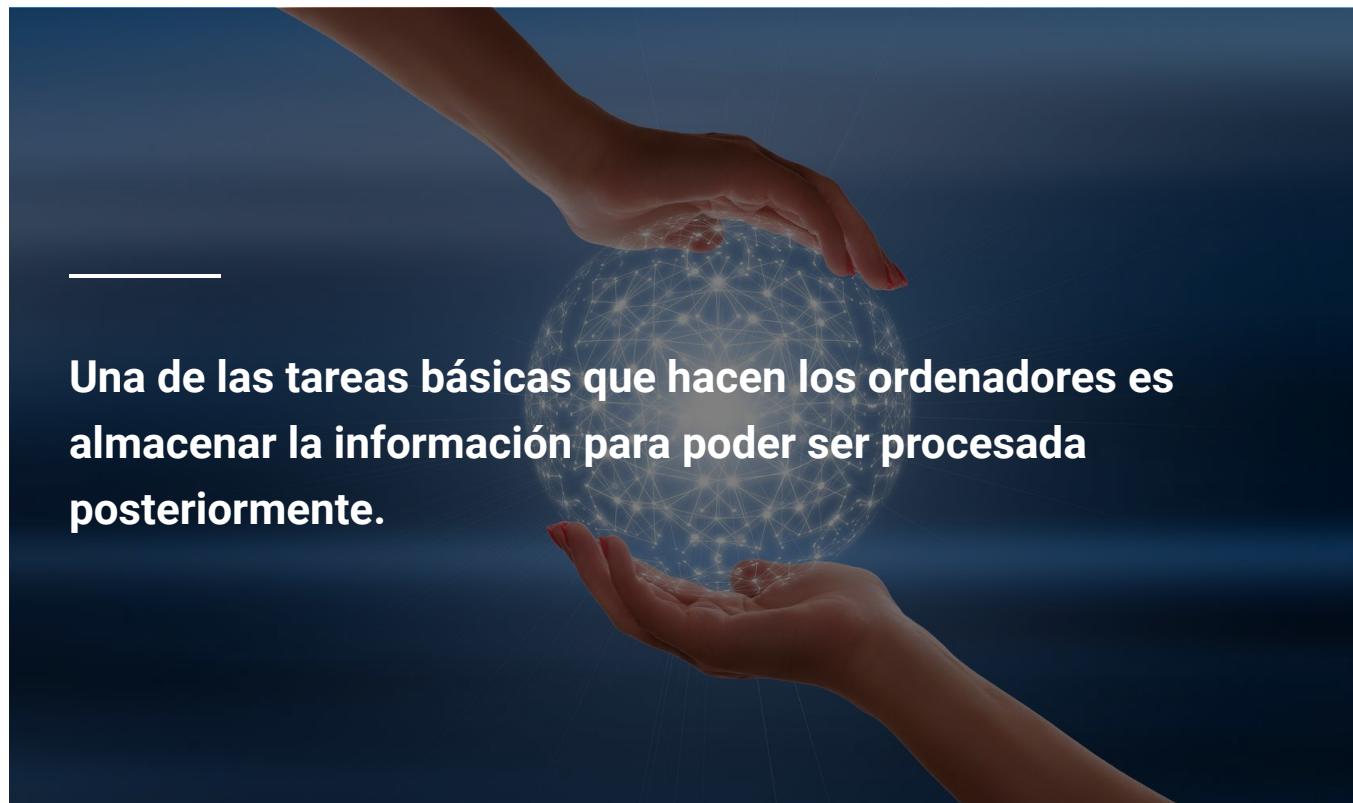
?

Autoevaluación

CONTENIDO PDF

☰ Contenido PDF

Intercambio de la información en la Web

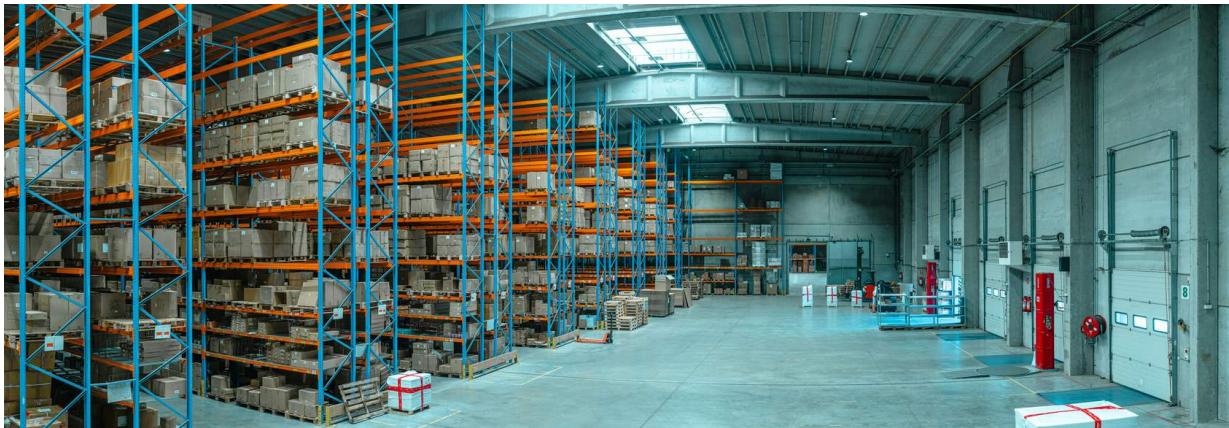


Una de las tareas básicas que hacen los ordenadores es almacenar la información para poder ser procesada posteriormente.

Introducción

La información puede ser de muchos tipos diferentes (texto, imágenes, vídeos, música ...) pero lo realmente importante será **cómo la almacena el ordenador** para poder tratarla

posteriormente de manera eficiente.



Ejemplo de almacenamiento

En el pasado, con los ordenadores centrales la información se generaba y se procesaba en el mismo lugar. Pero la aparición de los ordenadores personales, la eclosión de las redes y, sobre todo, el éxito de Internet, ha creado toda una serie de problemáticas que hasta el momento no existían.

Ahora los datos generados en un lugar ahora pueden ser consumidos en un lugar totalmente diferente, como:

- En sistemas operativos completamente diferentes.
- En máquinas que pueden funcionar de maneras muy diversas.

Por tanto, en un sistema informático moderno se debe tener en cuenta esta posibilidad a la hora de almacenar y sobre todo transmitir datos. En la actualidad son miles los servicios y aplicaciones que comparten información y de no ser por algunos estándares creados para representaciones de datos esta tarea sería un verdadero infierno. XML, CSV, JSON o YAML son algunos de esos formatos de intercambio de información.

XML

```
<Servers>
  <Server>
    <name>Server1</name>
    <owner>John</owner>
    <created>12232012</created>
    <status>active</status>
  </Server>
</Servers>
```

JSON

```
{
  Servers: [
    {
      name: Server1,
      owner: John,
      created: 12232012,
      status: active,
    }
  ]
}
```

YAML

```
Servers:  
  - name: Server1  
    owner: John  
    created: 12232012  
    status: active
```

En este [enlace](#) podrás consultar las diferencias entre JSON, XML y YAML

VIDEOTUTORIAL



LMSGI01 Introducción



LMSGI01 Introducción

Uploaded by asir edukiak on 2019-09-11.

VER EN YOUTUBE >

Introducción al lenguaje de marcas

El "lenguaje de marcas" se refiere a la forma de codificar un documento o página web.

¿Qué es un lenguaje de marcas?

Un lenguaje de marcas es **un modo de codificar un documento** donde, junto con el texto, se incorporan etiquetas, marcas o anotaciones con información adicional relativa a la estructura

del texto o su formato de presentación.

Los lenguajes de marcas permiten hacer explícita la estructura de un documento, su contenido semántico o cualquier otra información lingüística o extralingüística que se quiera hacer patente.

```
<tr class="row-1 row-first row-last">
  <td class="col-1 col-first">
    <div class="views-field views-field-title"> <span class="field-content"><a href="/yellowleather/products/shoes/comfy-leather-shoes">img type="image" alt="Yellow leather shoes product image" /</a></span></div>
    <div class="views-field views-field-uc-product-image">
      <div class="field-content">
        <a href="/yellowleather/products/shoes/comfy-leather-shoes">img type="image" alt="Yellow leather shoes product image" /</a>
      </div>
    </div>
    <div class="views-field views-field-display-price"> <span class="views-label views-label-display-price"></span></div>
  </td>
  <td class="col-2">
    <div class="views-field views-field-title"> <span class="field-content"><a href="/yellowleather/products/belts/embossed-spread-wing-eagle-belt">img type="image" alt="Yellow leather belts product image" /</a></span></div>
    <div class="views-field views-field-uc-product-image">
      <div class="field-content">
        <a href="/yellowleather/products/belts/embossed-spread-wing-eagle-belt">img type="image" alt="Yellow leather belts product image" /</a>
      </div>
    </div>
    <div class="views-field views-field-display-price"> <span class="views-label views-label-display-price"></span></div>
  </td>
  <td class="col-3 col-last">
    <div class="views-field views-field-title"> <span class="field-content"><a href="/yellowleather/products/hats/leather-hat">img type="image" alt="Yellow leather hats product image" /</a></span></div>
    <div class="views-field views-field-uc-product-image">
      <div class="field-content">
        <a href="/yellowleather/products/hats/leather-hat">img type="image" alt="Yellow leather hats product image" /</a>
      </div>
    </div>
    <div class="views-field views-field-display-price"> <span class="views-label views-label-display-price"></span></div>
  </td>
</tr>
```

Ejemplo de código HTML

El lenguaje de marcas más extendido es el **HTML** (HyperText Markup Language, lenguaje de marcado de hipertexto), fundamento del World Wide Web (entramado de comunicación de alcance mundial). Sin embargo, históricamente el marcado comenzó a usarse en la industria editorial y de la comunicación para definir el formato de los textos.

Algunos procesadores de texto optaron por guardar toda la información como texto, haciendo que las indicaciones de formato (tamaño de la página, color, tipografía, márgenes, etc.) también se almacenen de forma textual. Dichas indicaciones son caracteres marcados de manera especial para que así un programa adecuado pueda traducir dichos caracteres no como texto sino como operaciones que finalmente producirán mostrar el texto del documento de forma adecuada.



Información en una página web

La idea del marcado procede del inglés "*marking up*", término con el que se referían a la técnica de marcar manuscritos con lápiz de color para hacer anotaciones como por ejemplo la tipografía a emplear en las imprentas. Este mismo término se ha utilizado para los documentos de texto que contienen comandos u anotaciones y que han dado lugar a lenguajes llamados lenguajes de marcas, lenguajes de marcado o lenguajes de etiquetas.

El aspecto de un documento realizado con lenguaje de marcas sería el siguiente:

```
<carta>
    <fecha>22/11/2035</fecha>
    <presentacion>Estimado cliente:</presentacion>
    <contenido>bla bla bla ...</contenido>
    <firma>Don Jose Gutiérrez González</firma>
</carta>
```

Etiquetas o marcas



Los lenguajes de marcas utilizan una serie de etiquetas especiales o marcas intercaladas en un documento de texto sin formato.

¿Qué son las marcas?

Las **marcas** son una serie de **códigos** que se incorporan a los documentos electrónicos para determinar el formato, la forma en que se han de imprimir, la estructura de los datos, etc. Por lo

tanto, son anotaciones que se incorporan a los datos y que son posteriormente interpretadas por los intérpretes del lenguaje y que ayudan al procesado del documento.

Las marcas, por tanto, deben ser fácilmente distinguibles del texto normal (por su posición, por seguir una sintaxis especial, etc.). Las más usadas son las que están formadas por textos descriptivos y están rodeadas de los símbolos de "menor que" (<) y "mayor que" (>) y, normalmente, se utilizan dos etiquetas:

1

Etiqueta de inicio, la cual sirve para indicar que acaba de comenzar el efecto que pretendemos usar.

2

Etiqueta de fin, la cual utilizamos para señalar que ha terminado el efecto que queríamos presentar.

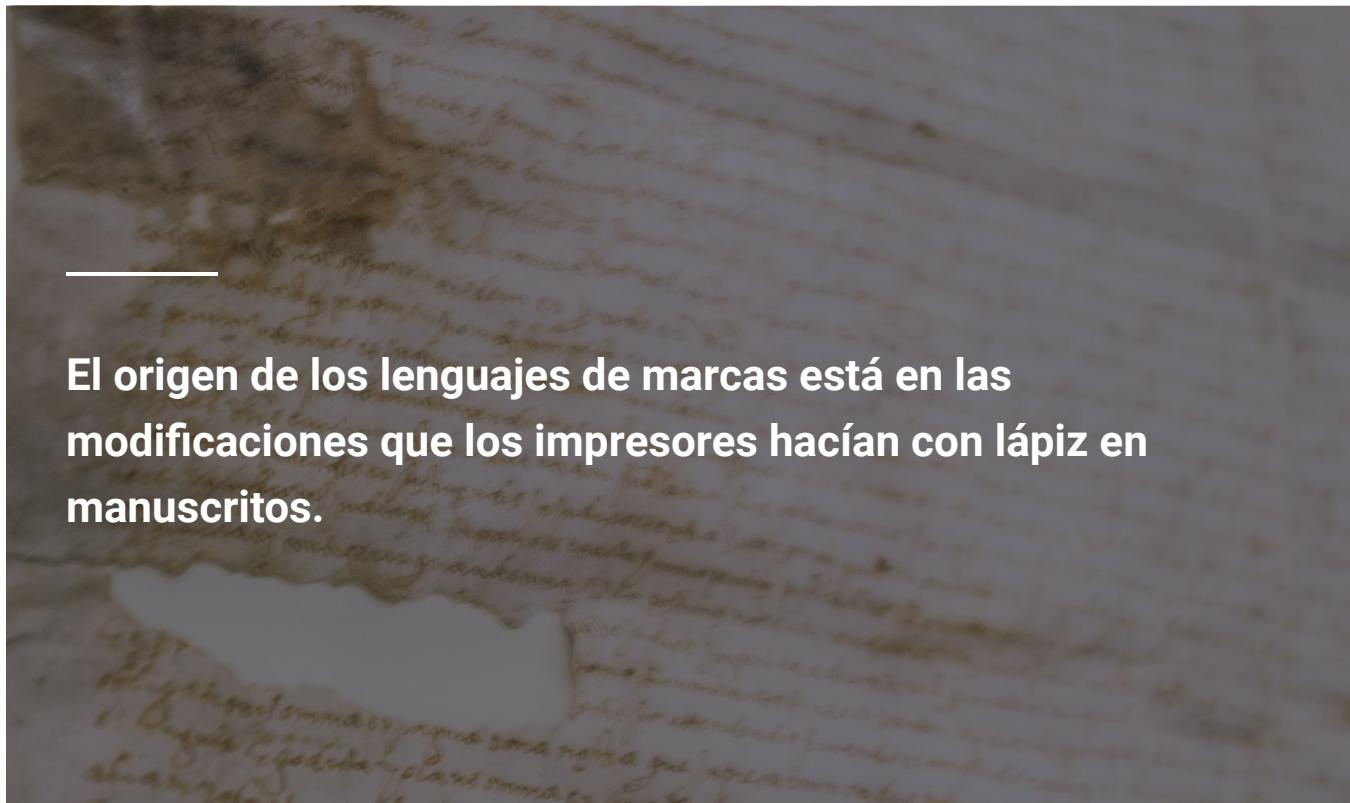
Por ejemplo, en el lenguaje HTML, este código:

```
<u>Esto está subrayado</u>
```

Un navegador lo interpreta así:

Esto está subrayado

Evolución histórica



El origen de los lenguajes de marcas está en las modificaciones que los impresores hacían con lápiz en manuscritos.

Los comienzos

Antiguamente, cuando alguien quería imprimir un libro que había escrito, los impresores, con un lápiz generalmente de color azul, escribían en el texto qué características debía tener cada

parte del texto, indicando si se tenía que hacer en negrita, si era el título del libro, etc. Se cree que estos son los antecedentes de las marcas.

En los años 70 con el desarrollo de los editores y procesadores de texto, surgen los primeros lenguajes informáticos especializados en tareas de descripción y estructuración de información: los lenguajes de marcas. Este marcado estaba exclusivamente orientado a la presentación de la información, aunque pronto se percataron de las posibilidades del marcado y le dieron nuevos usos que resolvían una gran variedad de necesidades, apareció el formato generalizado.

Evolución de los lenguajes de marcas

Analicemos la evolución de los lenguajes de marcas:

Años 70-80

SGML

Uno de los problemas que se conocen desde hace décadas en la informática es la **falta de estandarización en los formatos de información** usados por los distintos programas.

Para resolver este problema, en los años sesenta IBM encargó a Charles F. Goldfarb la construcción de un sistema de edición, almacenamiento y búsqueda de documentos legales. Tras analizar el

funcionamiento de la empresa llegaron a la conclusión de que para realizar un buen procesado informático de los documentos había que establecer un formato estándar para todos los documentos que se manejaban en la empresa. Con ello se lograba gestionar cualquier documento en cualquier departamento y con cualquier aplicación, sin tener en cuenta dónde ni con qué se generó el documento. Dicho formato tenía que ser válido para los distintos tipos de documentos legales que utilizaba la empresa, por tanto, debía ser flexible para que se pudiera ajustar a las distintas situaciones.

El formato de documentos que se creó como resultado de este trabajo fue **GML** (Generalized Markup Language), cuyo objetivo era describir los documentos de tal modo que el resultado fuese independiente de la plataforma y la aplicación utilizada.

El formato GML evolucionó hasta que en 1986 dio lugar al estándar ISO 8879 que se denominó **SGML** (Standard Generalized Markup Language). Éste era un lenguaje muy complejo y requería de unas herramientas de software caras. Por ello su uso ha quedado relegado a grandes aplicaciones industriales.

Imagen de Dreftymac, [Wikipedia](#)



```
<!DOCTYPE motd [ <!E
<motd>
<!-- created: 2003-12-12-->
<sentence>Do not throw
out the <keep>baby</>
with the
<refuse>dirty</>,
<refuse>stinky</>,
<refuse>bathwater</>.
</>
<!-- finish this later-->
</motd>
```

Años 90

HTML

En 1991, parecía que los editores WYSIWYG que almacenan los documentos en formatos binarios propietarios abarcarían casi la totalidad del procesamiento de textos, relegando al SGML a usos profesionales o industriales muy específicos. Sin embargo, la situación cambió drásticamente cuando Sir Tim Berners-Lee, que había aprendido SGML utilizó la sintaxis SGML para crear el **HTML**.

HTML es una **versión simplificada de SGML**, ya que sólo se utilizaban las instrucciones absolutamente imprescindibles. Además, era tan fácil de comprender que rápidamente tuvo gran aceptación logrando lo que no pudo SGML. HTML se convirtió en un estándar general para la creación de páginas web y su sencillez ha sido uno de los factores que ha llevado a la rápida popularidad del World Wide Web. El HTML es hoy en día el tipo de documento más empleado en el mundo.

Cualquier persona podía escribir documentos en este formato, sin apenas necesidad de conocimientos de informática. Esta fue una de las razones de su éxito, pero también condujo a un cierto caos. El crecimiento exponencial de la web en los años 90 produjo documentos en cantidades ingentes pero mal estructurados, problema agravado aún más por la falta de respeto por los estándares, por parte de diseñadores web y fabricantes de software.



Finales de los 90

XML

La respuesta a los problemas surgidos en torno al HTML vino de la mano del **XML** (eXtensible Markup Language). XML es un metalenguaje de etiquetas desarrollado por el World Wide Web Consortium (W3C) o Consorcio WWW, organización dedicada a la estandarización de casi todas las tecnologías ligadas a la web.

XML **no es un lenguaje** en el sentido de que no define un conjunto de etiquetas. XML **es un metalenguaje** porque permite crear lenguajes. Así, XML proporciona una serie de reglas para que cualquiera pueda definir su propio conjunto de etiquetas y atributos y pueda definir las relaciones que existen entre esas etiquetas.

Por lo tanto, tanto GML como SGML y XML, son metalenguajes. Los tres permiten crear lenguajes. Como hemos visto, SGML está inspirado por GML y a su vez, XML está inspirado por SGML.

Por otro lado, HTML, XHTML, MathML, RSS, SVG y otros más, sí que son lenguajes. HTML está basado en las reglas de SGML, mientras que los otros lenguajes que he nombrado están basados en XML.

Imagen de Dreftymac, [Wikimedia](#)

```
<?xml version="1.0"?>
<quiz>
  <qanda seq="1">
    <question>
      Who was the forty-second
      president of the U.S.A.?
    </question>
    <answer>
      William Jefferson Clinton
    </answer>
  </qanda>
  <!-- Note: We need to add
       more questions later.-->
</quiz>
```

XML

VIDEOTUTORIALES

EVOLUCIÓN HISTÓRICA

CASI TODO ES XML

 YOUTUBE

LMSGI01 Evolución Histórica



LMSGI01 Evolución Histórica

Uploaded by asir edukiak on 2019-09-11.

UNDEFINED YOUTUBE >

EVOLUCIÓN HISTÓRICA

CASI TODO ES XML

YOUTUBE

LMSGI01 Casi todo es XML



LMSGI01 Casi todo es XML

Uploaded by asir edukiak on 2019-09-11.

UNDEFINED YOUTUBE >

Características

```
<!DOCTYPE html PUBLIC "-//IETF//DTD HTML 2.0//EN">
<html dir="ltr" xmlns="http://www.w3.org/1999/xhtml">
<head>
    <title>Hello World</title>
    <meta http-equiv="Content-Type" content="text/html; charset=UTF-8"/>
    <meta name="keywords" content="Hello World, CSS, HTML, XML"/>
    <meta name="description" content="A simple Hello World page with CSS styling."/>
    <meta name="content-language" content="en"/>
    <link rel="stylesheet" type="text/css" href="style.css"/>
</head>
<body>
<div class="banner">
    <div style="margin:0 auto; width:300px; border:1px solid black; padding:5px; text-align:center; background-color:#f0f0f0; font-size:14px; font-weight:bold; color:#000000;">Hello World</div>
</div>
</body>
</html>
```

Los lenguajes de marcas se basan en texto plano sin formato.

Características

Entre las características propias de los lenguajes de marcas, destacan:

Se basan en texto plano

Los lenguajes de marcas se basan **en texto plano sin formato**. Estos caracteres pueden estar codificados en diferentes códigos de caracteres: ASCII, ISO-8859-1, UTF-8, etc.

Una de las ventajas que intentan aportar los lenguajes de marcas es que se pueden interpretar directamente independientemente de la plataforma, del sistema operativo o del programa que se use.

El hecho de que estén basados en formato de texto hace que sean fáciles de crear y modificar. Sólo requieren un simple editor de textos.

Permiten utilizar metadatos

En general, un grupo de metadatos se refiere a un **grupo de datos que describen el contenido informativo** de un objeto al que se denomina recurso. En este caso, las marcas son la forma en que se añaden los metadatos a los documentos de texto y cómo se consiguen superar las limitaciones del formato de texto.

Las marcas se intercalan entre el contenido del documento, por lo que generalmente estas etiquetas suelen ser descriptivos de qué es lo que indica el contenido de los datos que contienen.

Son fáciles de interpretar y procesar

Las diferentes etiquetas con distintos significados, permiten que posteriormente un programa pueda interpretar qué es lo que debe hacer con cada uno de los datos del documento.

Son fáciles de crear y suficientemente flexibles

A pesar de que fueron pensados para contener datos de texto, los lenguajes de marcas han demostrado que son capaces de contener **datos de muchos tipos diferentes**.

Actualmente se están utilizando archivos de marcas para representar imágenes vectoriales, fórmulas matemáticas, crear páginas web, ejecutar funciones remotas mediante servicios web, representar música o sonidos, etc.

Y sin importar qué tipo de datos se representen siempre habrá la posibilidad de crear estos archivos desde un editor de texto básico.

Herramientas básicas

Para trabajar con lenguajes de marcas es necesario editar los documentos y luego procesarlos, por tanto tenemos dos tipos de herramientas:

- **Editores.** Al basarse en ficheros de texto plano, basta utilizar un procesador de texto normal y corriente para construir un documento. Desde editores tan simples como el Bloc de notas de los sistemas Windows ó Gedit de sistemas Unix hasta editores más complejos como el Microsoft Word, pasando por editores especializados en XML como el Oxygen XML Editor.

- **Procesadores.** Para interpretar el código se puede utilizar cualquier navegador. Un procesador es un conjunto de módulos de software que comprueba que el documento cumple las normas establecidas para que pueda abrirse. En el caso de las páginas HTML, el procesador sería el

navegador que interpreta las etiquetas HTML y muestra el documento de manera adecuada.

Ejemplos

Algunos ejemplos de lenguajes de marcas son:

RTF	HTML	RSS	MATHML	

El Formato de Texto Enriquecido (**Rich Text Format**) es un formato de archivo informático desarrollado por Microsoft en 1987 para el intercambio de documentos multiplataforma. La mayoría de los procesadores de texto pueden leer y escribir documentos RTF.

RTF	HTML	RSS	MATHML	

El lenguaje de marcas de hipertexto (**Hypertext Markup Language**) hace referencia al lenguaje de marcado para la elaboración de páginas web.

RTF	HTML	RSS	MATHML	
------------	-------------	------------	---------------	--

La sindicación realmente simple (**Really Simple Syndication**) es un formato XML para distribuir contenido en la web que se actualiza frecuentemente a usuarios que se han suscrito a la fuente de contenidos.

RTF	HTML	RSS	MATHML	
------------	-------------	------------	---------------	--

Mathematical Markup Language es un lenguaje de marcado basado en XML, cuyo objetivo es expresar notación matemática de forma que distintas máquinas puedan entenderla, para su uso en combinación con XHTML en páginas web, y para intercambio de información entre programas de tipo matemático en general.

RTF	HTML	RSS	MATHML	
------------	-------------	------------	---------------	--

Gráficos vectoriales escalables, o gráficos vectoriales redimensionables (del inglés Scalable Vector Graphics) es un formato de gráficos vectoriales bidimensionales, tanto estáticos como animados, en formato XML, cuya especificación es un estándar abierto desarrollado por el W3C desde 1999.

RTF	HTML	RSS	MATHML	
------------	-------------	------------	---------------	--

Es un formato abierto de notación musical basado en XML. Fue diseñado para el intercambio de partituras, particularmente entre diferentes editores de partituras.

Autoevaluación

Pregunta

01/03

Los lenguajes de marcas se utilizan para:

- Dar formato a los documentos de texto.
- Definir la estructura de los datos de un documento.
- Permitir el intercambio de ficheros entre diferentes aplicaciones y plataformas.
- Todas las respuestas son correctas.

Pregunta

02/03

¿Cuál de las siguientes líneas es correcta?

- <i>Texto en cursiva
- <i>Texto en cursiva</i>
- <i>Texto en cursiva</i>
- <l>Texto en cursiva</l>

Pregunta

03/03

Para crear un documento HTML es necesario:

- Software especializado para la tecnología HTML.
- Herramientas de validación.
- Un block de notas y un navegador.
- Al menos, un editor HTML.

Contenido PDF



Documento descargable en PDF

Este documento contiene una versión en PDF de las lecciones anteriores. Ten en cuenta que los elementos interactivos no funcionarán de la misma manera en este formato, por lo que se recomienda complementar su uso con la versión interactiva.

Contenido PDF

Puedes visualizar y descargar el contenido PDF a través de este enlace

IR AL ENLACE