

Proyecto Lenguaje de Marcas

1ºDAW B.

Enlace de la presentación :

https://www.canva.com/design/DAFNWO6DII0/boavXDdwR3qSI-zpocliHg/view?utm_content=DAFNWO6DII0&utm_campaign=designshare&utm_medium=link&utm_source=homepage_design_menu

Docente:

García Blázquez, Olga.

Autores:

Callejas Rojas, Gonzalo.

Donoso Valle, José Juan.

García Téllez, Joaquín.

López López, María Auxiliadora.

Martínez Cortés, José Manuel.

Ramírez Torres, Yolanda.

Qué son y para qué sirven los lenguajes de marca.

Es un lenguaje que se enfoca en describir de manera estandarizada la forma en la que cada parte del contenido debe mostrarse mediante una serie de etiquetas, centradas principalmente, en gestionar opciones visuales como: la distancia en los márgenes, tamaño de letra, fuente, colores, subrayados, posición, espaciado, ... por lo que originalmente era usado por editores, autores e imprentas para maquetar obras.

Un ejemplo de lenguaje de marcado pero no informatizado serían las pautas escribir algo o un dictado:

<< Primera con mayúsculas, "buenas", coma, primera con mayúsculas y todo negritas, "fernando", cerramos negritas, pasa a la siguiente línea, deja sangría, primera letra en mayúsculas "le informo de que su" empezamos mayúsculas y negritas "deuda" cerramos mayúsculas y negritas "aun" empezamos negritas y mayúsculas "no" cerramos mayúsculas y negritas "ha sido pagada">> Y daría como resultado un correo electrónico con el texto:

“Buenas, **Fernando**

Le informo de que su **DEUDA** aún **NO** ha sido pagada ”

Hay quien confunde lenguajes de marca como HTML con un lenguaje de programación, pero nada más lejos de la realidad. Aunque ambos sean lenguajes informático no significa que tengan las mismas funciones, ya que mientras que mientras un lenguaje de programación tiene funciones lógicas, bucles, operaciones aritméticas, ... y características que le permiten interactuar y dar órdenes a un ordenador, los lenguajes de marca sólo señalan cosas visuales.

Actualmente el lenguaje de marcas más extendido se podría decir que es HTML, un lenguaje inventado hace más de 30 años en el CERN, por lo que ha marcado el crecimiento y evolución de la WWW desde su origen. Gran parte de su éxito es debido a que conseguía ser muy flexible sin sacrificar la simpleza. Parte del éxito de los lenguajes de marcas en la actualidad es que, debido a sus características internas es muy útil en la web semántica, ya que la búsqueda a lo largo de la página web es muy simple permitiendo indexarla en profundidad.

Introducción, contexto y orígenes.

Los lenguajes de marcado o marcas, provienen y toman su nombre de la práctica tradicional e histórica de marcar los manuscritos con instrucciones de impresión en los márgenes. Esta tarea, tan habitual en la industria editorial desde la aparición de la imprenta, fue dando lugar a un grupo de marcas estandarizadas, cuya esencia ha sido trasladada hoy al mundo de la informática.

En realidad, más que de lenguajes, podría hablarse de metalenguajes o conjuntos de reglas que se encargan de establecer y definir la forma digital de los documentos, bien para controlar su procesamiento, bien para representar su significado.

A pesar de existir proyectos anteriores, como el GenCode, la iniciativa que sentó las bases de los lenguajes de marcas actuales partió en los años 60 del investigador de IBM Charles F. Goldfarb, considerado hoy padre de este tipo de lenguajes por su participación en el desarrollo del Generalized Markup Language o GML.

El gran éxito del lenguaje GML propició su extensión e incluso adopción por parte del gobierno de los EE.UU., hecho que acrecentó la necesidad de estandarizarlo. Así, tras un largo proceso, en 1986, se convirtió en norma ISO (ISO 8879:1986) bajo el nombre de 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. De esta manera, a partir de la sintaxis de SGML se desarrolló en 1990 el Lenguaje de Marcas de Hipertexto o HTML (HyperText Markup Language) por Sir Tim Berners-Lee, creador de la WWW. En la actualidad, constituye la tecnología predominante en la construcción de páginas web.

Más adelante, en respuesta a las deficiencias surgidas en el entorno HTML se desarrolló en 1998, el lenguaje XML (eXtensible Markup Language), desarrollado y evolucionado y regulado por el Consorcio W3C (World Wide Web Consortium), un consorcio internacional que genera recomendaciones y estándares que aseguran el crecimiento de la World Wide Web a largo plazo.

XHTML surge en el año 2000, para expresar el lenguaje HTML como un lenguaje XML válido, y convertirlo de un lenguaje de marcas de procedimientos a un lenguaje de marcas descriptivo.

La versión definitiva de la quinta revisión del estándar HTML se publicó en octubre de 2014. Se trata de HTML5 y especifica dos variantes de sintaxis para HTML: una HTML conocida como HTML5, y una variante XHTML conocida sintaxis XHTML 5 que deberá servirse con sintaxis XML.

Esta es la primera vez que HTML y XHTML se han desarrollado en paralelo.

Podemos observar cómo la versión HTML5 y posteriores se han ido desarrollando centrándose más en la descripción del contenido y no en su presentación.

Clasificación de los lenguajes de marca.

Se suelen diferenciar tres clases de lenguajes de marcado, **Lenguajes orientados a presentación, procedurales y descriptivos**, aunque en la práctica pueden combinarse varias clases en un mismo documento.

Tipos de lenguaje de marca:

Lenguajes de presentación

Indica la apariencia, el formato de un texto. Es decir, indican cómo debe presentarse un texto (sin modificar su estructura, permite aumentar el tamaño de fuente, eliminar o añadir cursiva, subrayado, etc). Las etiquetas suelen estar ocultas al usuario por lo que es el software el que debe descifrar el código.

Las ventajas de este lenguaje reside en que resulta más fácil de elaborar que otros, sobre todo para cantidades pequeñas de información.

Las aplicaciones o programas que emplean este tipo de marcado son las aplicaciones de edición y los procesadores de texto. RTF, TeX, Docbook, Etc.

Ejemplo de lenguaje de presentación:

RTF (Rich Text Format - Formato de Texto Enriquecido)

```
{\rtf1\ansi{\fonttbl\font0\fswiss Helvetica;}\f0\pard
```

Esto es un texto en {\b negrita}.\par}

Lenguajes de procedimientos.

Orientado a la presentación igual que en el caso anterior, pero con la diferencia de que las etiquetas son visibles al usuario permitiéndole interactuar con ellas. Por esta razón, el programa que representa el documento debe interpretar estas etiquetas o marcas.

ejemplo:

- **HTML** (Hypertext Markup Language - Lenguaje de Marcado de Hipertexto): Su objetivo es la creación de documentos con un formato visual interesante para compartirlo en Internet.

```
<p>Esto es un texto con <strong>negrita</strong></p>
```

- **LaTeX**: Es un sistema de composición de textos, orientado a la creación de documentos escritos que presenten una alta calidad tipográfica.

```
\LaTeX{} es un programa para preparar documentos con el sistema de  
tipografías\footnote{%nota al pie de página Seg'un Wikipedia, la tipografía es el  
arte y técnica del manejo y selección de tipos, originalmente de plomo, para crear  
trabajos de impresión } %fin nota al pie de página\TeX{}.\newline
```

Lenguaje descriptivo o semántico.

Utilizan marcas o etiquetas para definir y describir las diferentes partes en las que se estructura el documento. Definen su contenido pero sin explicitar cómo ha de mostrarse.

- **XML** (eXtensible Markup Language - Lenguaje de Marcado Extensible): Es un metalenguaje que permite definir otros lenguajes de marcas. HTML provienen de este lenguaje.

```
<carta>
<fecha>22/11/2006</fecha>
<presentacion>Estimado cliente:</presentacion>
<contenido>bla bla bla bla ...</contenido>
<firma>Don José Gutiérrez González</firma>
</carta>
```

Ventajas y desventajas de los lenguajes de marcas.

1. Lenguaje de Presentación:

- **Ventajas:** Es útil para maquetar la presentación de un documento para su lectura, además resulta más fácil de elaborar, sobre todo para cantidades pequeñas de información.
- **Desventajas:** Resulta poco eficiente para el procesamiento automático de la información y para mantenerla o modificarla, por lo que su uso se ha ido reduciendo en proyectos grandes.

2. Lenguaje de Procedimientos:

- **Ventajas:** Las etiquetas son visibles para el usuario, lo que permite que el mismo pueda interactuar con ellas.
- **Desventajas:** Este tipo de marcado se ha usado extensivamente en aplicaciones de edición profesional, manipulados por tipógrafos calificados, ya que puede llegar a ser extremadamente complejo.

3. Lenguaje Descriptivo o semántico:

- **Ventajas:** Dada su simplicidad para estructurar textos y capacidad para establecer enlaces con otros archivos es, actualmente, la tecnología dominante en la-

construcción de páginas web. Los fragmentos de texto se etiquetan tal como son, y no tal como deben aparecer por lo que es un lenguaje bastante flexible.

- **Desventajas:** Utiliza etiquetas para describir los fragmentos de texto, pero sin especificar cómo deben ser representados, o en qué orden.

Cuáles son los lenguajes de marca más usados.

RTF -> Rich Text Format (formato de texto enriquecido), es un formato de archivo informático desarrollado por Microsoft en 1987 para el intercambio de documentos multiplataforma.

```
{\rtf1\ansi{\fonttbl{\f0\fswiss Helvetica;}\f0\pard
  Esto es un texto en {\b negrita}.\par
}
```

Se muestra así cuando es representado por un programa que soporta RTF:

Esto es un texto en negrita.

HTML -> Es el componente más básico de la Web. Define el significado y la estructura del contenido web, se distingue de otro texto en un documento mediante "etiquetas", que consisten en el nombre del elemento rodeado por "<" y ">". El nombre de un elemento dentro de una etiqueta no distingue entre mayúsculas y minúsculas. Es decir, se puede escribir en mayúsculas, minúsculas o una mezcla.

Este texto va en negritas.

El resultado será:

Este texto va en negritas.

XML -> Son las siglas de Extensible Markup Language, Normalmente, los archivos XML contienen información de cualquier tipo, ya sea información sensible o información de cualquier ámbito.

Los archivos XML se componen de etiquetas que nos aportan datos e información que queremos procesar. Estas etiquetas pueden estar de forma individual o anidadas.

Habitualmente un fichero XML incluye mucha información y debe de ser procesada correctamente por el usuario, en este caso el desarrollador. Cuanto más grande sea un fichero XML nos estará indicando que más información trae.

Lo primero que debemos saber de un fichero XML es que siempre contendrá una única etiqueta dando la introducción a dicho fichero, y a partir de ella podremos crear cualquier tipo de archivo XML.

Dentro de estos archivos XML tendremos una o varias etiquetas, que a su vez tendrán otras etiquetas, que podrán estar o no anidadas.

```
<message>
  <warning>
    Hola, mundo
  </warning>
</message>
```

Postscript -> Es un [lenguaje de descripción de páginas](#), utilizado en muchas [impresoras](#) y también muy común como formato de transporte de archivos gráficos en talleres de impresión profesional.

Un ejemplo de una pila de instrucciones para realizar un dibujo simple con PostScript se detalla a continuación:

```
newpath
% Inicialización del cursor
100 100 moveto
% Dibujo del rectángulo
300 100 lineto
300 250 lineto
100 250 lineto
100 100 lineto
stroke
```

- *newpath* indica que se va a dibujar una nueva figura.
- *moveto* sitúa el cursor para comenzar a dibujar la figura. En este caso dichas coordenadas son 100, 100.
- Los distintos *lineto* dibujan las líneas de la figura (en este caso todas son líneas rectas).
- Finalmente, el comando *stroke* indica la realización del dibujo al intérprete.

Bibliografía.

COOMBS James H., RENEAR, Allen H., DE ROSE, Steven J. (1995). *"Markup Systems and the Future of Scholarly Text Processing"*. Detroit, Estados Unidos: Universidad de Michigan

FERNÁNDEZ-VALMAYOR, A., NAVARRO, A., FERNÁNDEZ-MANJÓN, B. y SIERRA, J. L. *"Lenguajes de programación, lenguajes de marcado y modelos hipermedia: una visión interesada de la evolución de los lenguajes informáticos"*. Madrid: Universidad Complutense de Madrid.

GARCÍA CASTAÑO Carolina y ARROYO MENÉNDEZA David (2002). *"Biblioteca Digital y Web Semántica"* Madrid.

GOLDFARB, Charles F. (1996). *"The Roots of SGML -- A Personal Recollection"*. Oxford University Press.

GOLDFARB, Charles F. (1990). *"The SGML Handbook"*. Oxford University Press.

MORALES, Rafael (2018). *"ticARTE" — Qué son los lenguajes de marca.*

MORENO PÉREZ, Juan Carlos y GONZÁLEZ RUIZ Sergio Luis.. *"Lenguajes de marcas y sistemas de gestión de información"*. Madrid.

TABOADA LEÓN, Juan José (2013-2022). *"Docencia Juan José Taboada León"* – Módulo: Lenguajes de marcas y sistemas de gestión de la información. Huelva.

Enlaces bibliográficos:

<https://www.abrirlave.com/lmsgi/libro.php> [Consultado el 22/09/2022]

<https://developer.mozilla.org/es/docs/Web/HTML> [Consultado el 22/09/2022]

<https://www.docencia.taboadaleon.es> [Consultado el 26/09/2022]

<https://es.slideshare.net/MariaHanse/lenguaje-de-marcas> [Consultado el 26/09/2022]

<https://www.ticarte.com/contenido/que-son-los-lenguajes-de-marcas> [Consultado el 24/09/2022]

<https://trabajolms.wordpress.com/2015/12/06/4-5-ventajas-y-desventajas-de-los-lenguajes-de-marcas/> [Consultado el 23/09/2022]

<http://w3.org/> [Consultado el 26/09/2022]

<https://developer.mozilla.org/es/docs/Web/HTML> [Consultado el 26/09/2022]

<https://home.web.cern.ch/science/computing/birth-web> [Consultado el 26/09/2022]

<https://www.ticarte.com/contenido/que-son-los-lenguajes-de-marcas> [Consultado el 26/09/2022]