

Lenguajes de marcas y Sistemas de gestión de la información

DEFINICIÓN Y TIPOS DE LENGUAJES DE MARCAS

DEFINICIÓN DE LENGUAJE DE MARCAS

Los **lenguajes de marcas**, también denominados lenguajes de marcado o lenguajes de etiquetas, son aquellos que complementan la información de un documento con marcas o anotaciones que modifican la estructura del documento o la forma de representarlo.

Dicho lenguaje de marcas especifica cuáles son las etiquetas posibles, el significado de cada una de ellas y dónde deben colocarse.

Dichas etiquetas no son visibles por el usuario, solo es visible el contenido y apariencia final del documento.

EJEMPLOS DE LENGUAJE DE MARCAS

- **LaTeX**, que permite crear documentos con expresiones científicas y fórmulas matemáticas, utilizando una tipografía y un lenguaje común para todas las computadoras.
- **RTF o Rich Text Format**, desarrollado por Microsoft para producir documentos de texto que incluyan anotaciones de formato.
- **SGML**, es el estándar de lenguajes de marcado que define, entre otros, que las etiquetas deben estar delimitadas por los símbolos < y >, y se cierran utilizando el símbolo /. Se trata del padre de XML y la base sobre la que se sostiene HTML.
- **HTML o Hypertext Markup Language**, que permite crear documentos transportables a través de Internet, en los que es posible marcar texto para crear enlaces entre documentos.
- **XML o eXtended Markup Language**, se trata de un subconjunto de SGML que permite definir lenguajes de marcado con sintaxis estricta pero extensible.

TIPOS DE LENGUAJE DE MARCAS

Los lenguajes de marcas se suelen clasificar en tres tipos, aunque existe la posibilidad de que un lenguaje de marcas combine las características de más de un grupo.

- **Orientados a la presentación**, se añaden etiquetas al texto del documento que contienen indicaciones del formato en que debe presentarse dicho texto al usuario. Sería el caso de los procesadores de texto como Microsoft Word y del lenguaje HTML.
- **Orientados a la descripción**, en los que las etiquetas dan un significado al texto, pero no indican cómo se debe presentar al usuario. Sería el caso de XML.
- **Orientados a procedimientos**, en los que las etiquetas se interpretan como órdenes a seguir, de modo que el documento en realidad contiene instrucciones que indican qué hacer con el texto del documento. Sería el caso de LaTeX y Postscript.

ORGANIZACIONES DESARROLLADORAS

Las organizaciones que se encargan de desarrollar y normalizar los lenguajes de marcas son los siguientes:

- **ISO o International Organization for Standardization**, que se creó al final de la Segunda Guerra Mundial y se encarga de promover el desarrollo de normas internacionales de fabricación, comercio y comunicación para todas las ramas industriales, a excepción de la eléctrica y electrónica.
- **W3C o World Wide Web Consortium**, que se creó en 1994 por el MIT y se encarga de tutelar el crecimiento y organización de la web. Su primer trabajo consistió en normalizar el lenguaje HTML. Posteriormente, desarrolló el estándar XML que fue publicado en 1998.

ETIQUETAS, ELEMENTOS Y ATRIBUTOS

Las partes de un documento de lenguajes de marcas se clasifican en:

- **Etiqueta** o **Tag**, es un texto delimitado por los símbolos < y >. Se distinguen entre etiquetas de inicio y etiquetas de fin. Las etiquetas de fin son como las de inicio añadiendo el símbolo /.

<nombre>María</nombre>

- **Atributo**, es un par nombre-valor que se encuentra dentro de la etiqueta de inicio y define los valores para las distintas propiedades que puede llevar asociada una etiqueta.

<ciudad provincia="Madrid">Alcobendas</ciudad>

- **Elemento**, que representa una estructura organizativa del documento, está formada por una etiqueta de inicio, una etiqueta de fin y todo lo que se encuentra entre ambas. Algunos elementos pueden estar vacíos y en ese caso no llevan etiqueta de fin.

GRAMÁTICAS

Todos los documentos de un mismo lenguaje de marcas tienen en común una gramática, que define las combinaciones de símbolos sintácticamente correctas.

Las gramáticas se definen mediante el **DTD** o **Document Type Definition**, que contiene las reglas de los elementos necesarios en la elaboración de un documento, como el nombre, su significado, dónde pueden ser utilizados y qué pueden contener.

XMLSchema o **XSD** es la evolución de DTD descrita por el W3C. Se trata de un lenguaje de definición más complejo y potente que utiliza sintaxis XML. Permite definir especificaciones de forma más detallada, pero consume más recursos y tiempo debido a su complejidad.