XML es una simplificación de SGML que elimina muchas de sus características más complejas, con el objeto de facilitar la creación de procesadores capaces de tratar los documentos escritos en alguno de sus lenguajes.

Estas simplificaciones se concretan en:

* Se define que las marcas sólo pueden expresarse del modo predeterminado, o sea, <marca> para apertura, y </marca> para cierre.
* Toda etiqueta debe ser expresamente abierta y cerrada.
* Los atributos de las marcas siempre requieren un valor y que éste se entrecomille.
* Se define también una única notación para definir entidades.
* Los nombres son sensibles al uso de mayúsculas y minúsculas.

XML desde su aparición ha sido ampliamente adoptado, aunque también ha recibido críticas

Entre sus principales aplicaciones está:

* La creación de **documentos semánticos estructurados**
* Como formato para el **intercambio de datos**
* Para el almacenamiento de cantidades moderado de datos estructurados

Anatomía del documento

Los documentos XML están constituidos por un conjunto de componentes de los que es imprescindible conocer su nombre y qué refieren:

Instrucción de procesamiento

Etiqueta

Texto

Elemento

Atributo

Comentario

Entidad

Nodo

#### **Estructura jerárquica**

Tanto en SGML como en XML los elementos debes estar perfectamente anidados unos dentro de otros, esto es, dado un elemento siempre habrá un único nodo que lo contenga completamente, excepto el caso del elemento raíz que es el que abraca todo el contenido y no puede estar incluido dentro de nada