

## XML: BREVE HISTORIA DE UN LENGUAJE DE MARCADO

XML es un lenguaje informático que se describe como metalenguaje permitiendo definir lenguajes de marcas para almacenar datos de forma legible, lo que significa que su diseño está destinado a operar prácticamente en cualquier sistema informático y facilitando la comunicación entre diferentes plataformas, pudiendo estructurar grandes volúmenes de datos y ser procesados.

El origen de XML (abreviatura de eXtended Markup Language) se sitúa en 1996. Fue un proyecto de la W3C, una organización dedicada a la implantación de estándares para el desarrollo de Web (similar a la labor de la organización ISO). Para entender este lenguaje primero debemos volver la vista atrás: su origen se remonta a la década de 1970, tiempo en el que la mundialmente famosa IBM desarrolló GML debido a que la empresa necesitaba almacenar una ingente cantidad de información y manipularla entre diferentes sistemas operativos y plataformas. La ISO adoptó este sistema de información y trabajó en una versión estandarizada que se denominó SGML, el cual vio la luz en 1986 y fue el germen de multitud de sistemas de gestión de información posteriores.

El siguiente paso fue la creación de la World Wide Web en 1989, obra del ingeniero informático Tim Berners Lee. Este faraónico proyecto requirió de un tipo de lenguaje acorde a las necesidades de dicho producto, y la creación de HTML en 1992 por parte de la W3C (organización fundada por Berners) fue, dentro del marco de SGML, el estándar más extendido que impulsó la expansión de Web. HTML establecía un determinado DTD (Definition Type Document) que describe una estructura y sintaxis mediante etiquetas para la creación de estructuras comunes de datos fáciles de entender tanto por máquinas como por usuarios para que dichos documentos puedan ser validados.

Pero HTML presentaba limitaciones, como que cada documento tiene un vocabulario fijo establecido por cada DTD, lo que impide usar combinaciones de varios vocabularios dentro de un mismo documento. Además, si un navegador no dispone del vocabulario empleado para la creación de una página web, éste no podrá analizar e interpretar dicha página. Otra limitación importante es que cada nueva versión de HTML suprime algunas características de las versiones anteriores modificando su código, lo que

obliga a las páginas web creadas con una cierta versión anterior de HTML a actualizarse con los nuevos cambios de la última versión del lenguaje o si no presentarán problemas de visualización. Esto hace que las páginas web sin revisión periódica sean susceptibles de quedar obsoletas y obliga a los navegadores web a ser cada vez más complejos para incorporar en su programación versiones compatibles con versiones anteriores de HTML para poder interpretar páginas web cuyo código no ha sido actualizado a la nueva versión de HTML.

XML se ideó tomando ciertos aspectos de SGML y HTML y desechando otros para facilitar la interpretación de documentos por parte de programas automáticos. Aunque esta era un principio la idea inicial del desarrollo de XML, con el tiempo fue derivando al intercambio de datos altamente estructurados entre aplicaciones, lo que le provocó que le llovieran críticas por el nivel de complejidad requerido en su uso.

Sin embargo, pese a que lo más destacable sobre la evolución de los lenguajes de marcado fue el intento de adopción de un estándar que fijase la norma a seguir, en la práctica tanto XML como los lenguajes anteriores se desarrollaron por caminos paralelos: mientras que la W3C como la ISO desarrollaban un estándar, otras versiones alternativas de estos lenguajes se desarrollaban a la par por diversas comunidades de programadores libres como JSON y YAML, los cuales se presentan como alternativas a XML ya que se centran en la representación de datos estructurados en lugar de documentos narrativos.