

# LENGUAJES DE MARCAS Y SISTEMAS DE GESTIÓN DE INFORMACIÓN

XML: UD1, Introducción y un poco de Hª

# Índice

2

- Introducción
- Un poco de historia
- GML, SGML, HTML, XML

# XML: antes de..

3

- Es libre y abierto.
- No es tan nuevo: Viene de años 70
- Ahora en plena evolución o caída..
- Documentos:
  - antes texto=documento.
  - Ahora: documento es un fichero electrónico para cualquier media (texto, video, audio..)
- Destinatarios del documento:
  - Personas que los “consumen”
  - Máquinas que los “consumen”, para presentarlos o combinarlos con otros, etc.

# Concepto de “marca”

4

- Forma de añadir información sobre los datos junto a los propios datos.
- Insertamos símbolos o caracteres en un documento para indicar su estructura lógica o física
- Objetivo: separar los datos de su posterior tratamiento (por ejemplo apariencia)
- Si has estudiado HTML, que también es un lenguaje de marcas te será familiar.

# XML

5

- **NO** es un lenguaje de programación
  - Es un lenguaje de marcas
  - Es un “metalenguaje”
- Un metalenguaje es un lenguaje usado para describir formalmente otro lenguaje.
- XML permite a los desarrolladores crear lenguajes especializados,
  - todos en común: respetan la especificación XML.

# Historia (I)

6

- Ya existía algo en los 60s..
  - documentos con códigos de control para definir presentación del texto.
  - Eran formatos **propietarios, cerrados**, dependientes del programa que procesaba.
- Finales de los 60 algunos pedían
  - Formato estándar y común para publicaciones (p.e. libros)
  - Separar el contenido, la información, de la forma en la que se va a presentar

# Historia (I)

7

- GML de IBM, año 1971: Generalized Markup Language
  - Extensible
  - Deberían existir mecanismos para comprobar si eran válidos
  - Deberían ser “legibles” a las personas.
- Otras tendencias
  - DTP aplicaciones: TEX, LaTeX, etc. Parecido, pero orientados a presentación.

# Historia (II): Ejemplo GML

8

**:h2.** Definitions:

**:ol.**

**:li.1.** noun, a gem variety of corundum in transparent or translucent crystals of a color other than red; especially, a transparent rich blue

**:li.2.** noun, a gem of such corundum

**:li.3.** noun, a deep purplish blue color

**:li.4.** adjective, made of or resembling a sapphire gem

**:li.5.** adjective, having the color of a blue sapphire

**:eol.**



# SGML: *Standard Generalized Markup Language*

9

- ANSI (American National Standards Institute) a partir de GML 1980-84.
  - ISO(International Organization for Standardization): 1986, con ANSI
- Se centra relación estructura y contenido. Separa
  - Estructura de los datos
  - Contenidos
  - Estilo
- Formaliza concepto DTD: documento separado que especifica la estructura del mismo.

# ¿y por qué no SGML?

10

- Es un estándar enorme, muy complejo
- Los lenguajes SGML son enormes y consumen demasiados recursos.
- navegadores no lo soportan.
- Por ejemplo, añadir SGML a un procesador de texto dobla o triplica el precio del mismo..

# HTML: antes de XML

11

- HyperText Markup Language (HTML)
  - desarrollado en CERN (European Organization for Nuclear Research).
- Propósito original: facilitar mecanismos independientes de plataforma para documentos hipertexto,
  - con enlaces entre ellos, ya sea en el mismo servidor, red de área local o Internet.
- Internet ya era tecnología madura, soportada por empresas, universidades, gobiernos
- HTML fue de gran ayuda para difusión de contenidos de forma sencilla y rápida.

# HTML

12

- “Gratis”, simple, muy difundido y soportado
  - => se hizo muy popular.
  - => casi cualquiera podía tener una página web.
- Estas ventajas lo hacen al mismo tiempo deficiente:
  - no se puede extender, no hay estructura semántica, las empresas añaden etiquetas “propietarias”, etc.
  - La mayoría de los diseñadores de páginas web solo se preocupaban de la apariencia del documento.

# Vamos a arreglarlo..

13

- 1996: W3C (World Wide Web Consortium) patrocina grupo expertos en SGML
  - desarrollar lenguaje de marcas con la **potencia** y extensibilidad de **SGML**, pero con la **simplicidad** de **HTML**.
- Eliminaron todo aquello que no fuera básico, que no se usa, etc.
  - => lenguaje más pequeño, fácil y simple de implementar, orientado a la Web
- Así nace XML 1.0 alrededor de 1998.

# Definición en la norma

*"The Extensible Markup Language (XML) is a subset of the Standard Generalized Markup Language. . .whose goal is to enable generic SGML to be served, received, and processed on the Web in the way that is now possible with HTML. XML has been designed for ease of implementation and for interoperability with both SGML and HTML."*

# XML

15

- Es un lenguaje de marcas y un metalenguaje
  - los desarrolladores crean lenguajes específicos con sus propias etiquetas, estructuras, etc.
- No limita marcas a crear, sólo impone restricciones.
- Independiente del SW y de la plataforma
  - un mismo documento XML puede ser usado en cualquier SO con cualquier herramienta de desarrollo o publicación.
- Es un formato público, un estándar abierto. No depende de empresa o país.
- Fácil de implementar: herramientas gratuitas.

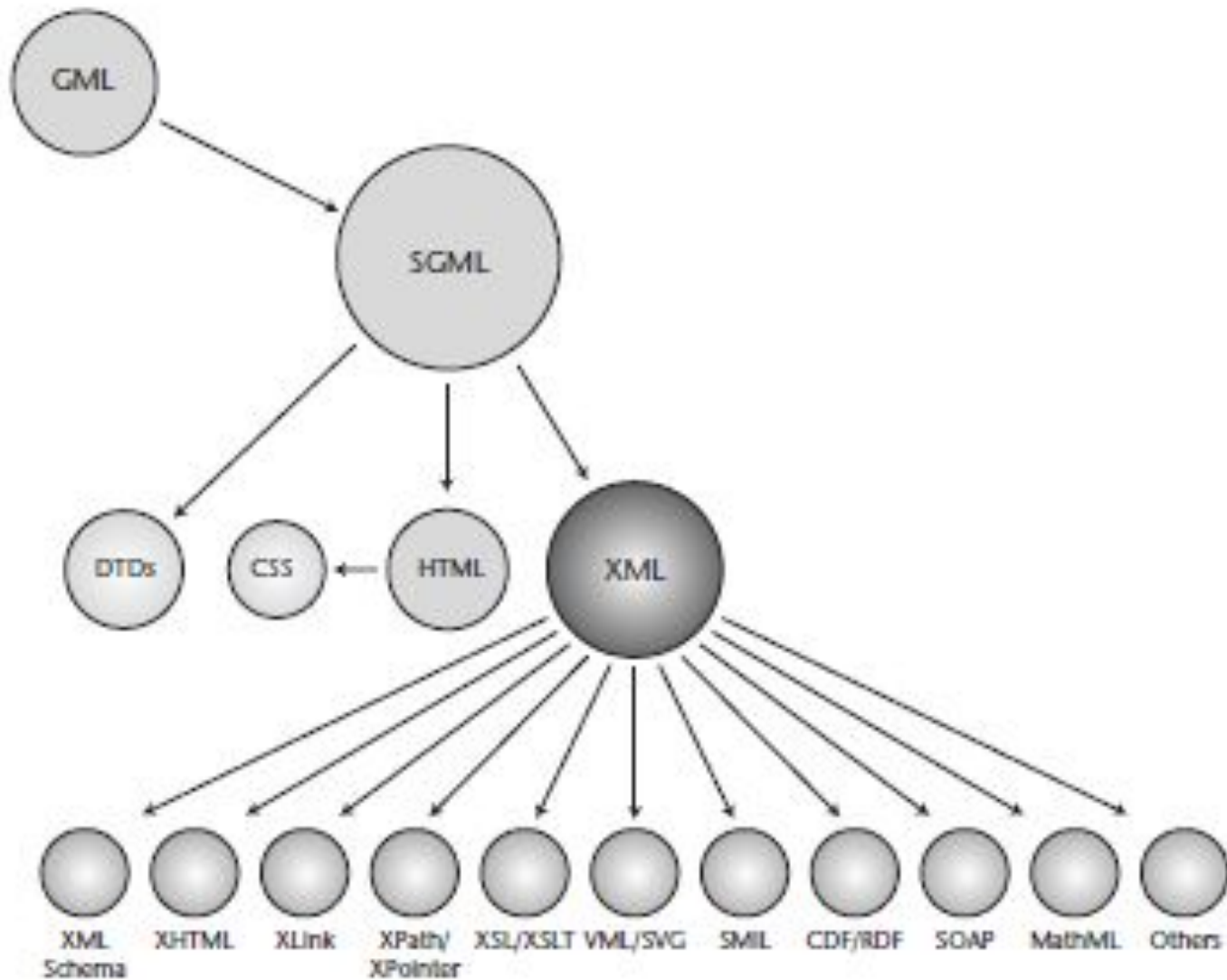
- El **diseñador** se puede **centrar** en uso de etiquetas, en **contenidos** y su **estructura**.
  - Formatear el contenido es un proceso posterior (e incluso a veces secundario)
- A los ficheros xml
  - se les comprueba la sintaxis,
  - y se puede comprobar la estructura (validación)



- Pensar en XML como en una versión menor de SGML, no como HTLM++
  - XML se diseñó pensando en la estructura y significado de los datos.
  - HTML, en cambio, con la forma en la que se presentan los datos.
- Las etiquetas XML describen datos y se usan como parte de transacciones de negocios, intercambios de información, etc.
  - la sintaxis XML requiere precisión y seguir estrictamente las reglas gramaticales.
  - Esto no ocurría con HTML hasta hace poco.

# Gráficamente

18



# Ejercicios

19

- XML en el PC:
  - buscar ficheros con extensión xml
  - [Configuración Filezilla](#)
- Examinarlos
  - Abrirlos con editor de texto o navegador
- XML en la web: [www.w3.org](http://www.w3.org)
- Otros ejemplos de uso de XML
  - [Docbook](#)
  - [SVG](#)
  - [Android](#)
  - ...
- Otro lenguaje de marcas: markdown