
La información, los lenguajes de marcado y su aplicación a la interacción hombre-computadora

Baltasar Fernández Manjón



Dpto. de Sistemas Informáticos y Programación,
Universidad Complutense de Madrid
Avda. Complutense s/n. 28040, Madrid, Spain,
Phone: (+34) 91 394 43 83.

e-mail: baltasar@sisp.ucm.es
<http://baltasar.sisp.ucm.es/~baltasar>

La información, los lenguajes de marcado y su aplicación a la interacción hombre-computadora

1. Motivación: World Wide Web
2. Lenguajes de marcado
3. Lenguaje de marcado generalizado estándar (SGML)
4. Estándares relacionados
5. Aplicaciones en interacción hombre-computador (HCI)

¿Éxito de Internet y del World Wide Web?

- m Web: un nuevo medio de comunicación de masas
 - o Revolución en la distribución, acceso e interacción con la información electrónica
 - o Información hipermedia: multimedia + hipertexto
- m World Wide Web
 - o Conjunto de protocolos de comunicación de datos
 - HTTP, FTP, Gopher, etc
 - o Sistema de nombrado y localización de recursos (URLs)
 - o Formato común de intercambio de datos
 - Lenguaje de marcado de hipertexto (HTML)
- m Utilización de un lenguaje de marcado (HTML)
 - o Compartición de información entre sistemas heterogéneos
 - o Enlazamiento de esa información

©Baltasar Fernández Manóñ

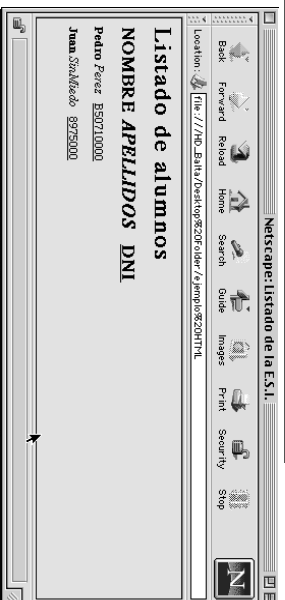
Lenguaje de marcado de hipertexto (HTML)

- m Separación entre representación y procesamiento
 - o Codificación del contenido y estructura de un documento median te un lenguaje estándar de marcas
 - o El navegador realiza y decide la representación de la información mediante el procesamiento de dichas marcas
- m Limitaciones
 - o Orientado a la representación de información
 - o No es extensible ni adaptable (conjunto fijo de marcas)
- m HTML es una aplicación del Lenguaje de marcado generalizado estándar (SGML)
- m Futuro HTML
 - o XML - Lenguaje de marcado extensible

©Baltasar Fernández Manóñ

Ejemplo HTML

```
<HTML>
<HEAD>
<TITLE>Listado de la E.S.I.</TITLE>
</HEAD>
<BODY>
<H1> Listado de alumnos </H1>
<H2> <B>NO MBRE</B> <B>APELLIDOS</B> <B>DNI</B> </H2>
<P> <B>Pedro</B> <B>Perez</B> <B>B50710000</B>
<P> <B>Juan</B> <B>Simi edo</B> <B>8975000</B>
</BODY>
</HTML>
```



©Baltasar Escudélez Marañón

Lenguajes de marcado generalizado

m Técnica descriptiva para representar documentos electrónicos simplificando su procesamiento automático

m Marcado generalizado

- o Descripción, mediante marcas, de la estructura y otras características del documento
- Metainformación: Estructura, organización, propósito
- Utilización de la misma información con diferentes finalidades
- o Técnica rigurosa para permitir procesamiento automático

m Documento = Contenido + Marcas

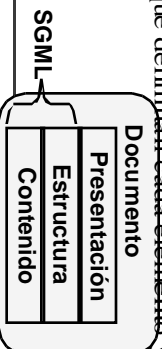
m Maximizar la usabilidad de la información

- o Accesibilidad, manteniendo la información

o Distintos procesamiento, independencia de plataforma
©Baltasar Escudélez Marañón

Lenguaje de marcado generalizado estándar (SGML)

- Metalenguaje estándar para la definición de lenguajes de marcado (ISO 8879, 1986)
- Método formal y general para diseñar clases de documentos estructurados para cada necesidad concreta
 - Estructuración del documento dependiente del contenido, significado o utilización de los datos
 - Especifica modelos jerárquicos de documentos
 - Características y composición de cada elemento del documento
 - Sintaxis de marcado para definir los elementos
 - Marcas que delimitan cada elemento y atributos aplicables



©Baltasar Escandell Marañón

Definición del tipo de documento (DTD)

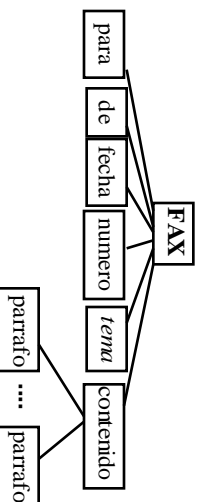
- Cada clase genérica de documentos se representa mediante su *Definición de Tipo de Documento (DTD)*
 - Con junto de reglas que especifican cual es la estructura lógica del documento como una jerarquía de elementos marcados (árbol)
 - Orden de elementos, su obligatoriedad, número de ocurrencia, marcas delimitadoras, tipo de contenido
 - Atributos
 - Información complementaria sobre los elementos
- Validación automática de documentos
 - Dado un documento se puede verificar automáticamente si cumple el conjunto de reglas especificado en su DTD
- SGML no se ocupa de la presentación de los documentos
 - La presentación es sólo un tipo más de procesamiento

©Baltasar Escandell Marañón

Ejemplo: definición de tipo de documentos FAX

Fax DTD

```
<!DOCTYPE fax I
<ELEMENT fax -- (para, de, fecha, numero, tema?, contenido)>
<ELEMENT para -- (#PCDATA)>
<ELEMENT de -- (#PCDATA)>
<ELEMENT fecha -- (#PCDATA)>
<ELEMENT numero -- (#PCDATA)>
<ELEMENT tema -- (#PCDATA)>
<ELEMENT contenido -- (parrafo+)>
<ELEMENT parrafo -- (#PCDATA)>
|>
```



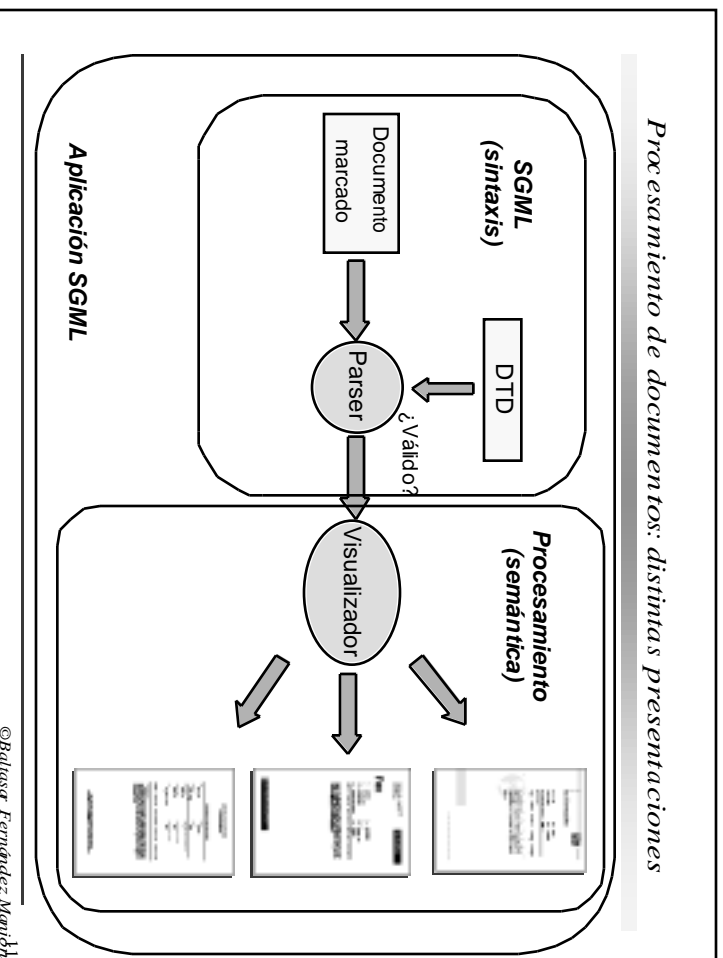
©Baltasar Fernández Manóñ

FAX: documento marcado

```
<!DOCTYPE fax SYSTEM "fax.dtd">
<fax>
  <para> Mmanuel Ortega </para>
  <de> Baltasar Fernández </de>
  <fecha> 16/9/99 </fecha>
  <numero> 92629 53 00 </numero>
  <tema> Curso Interacción Persona-Computador </tema>
  <contenido>
    <parrafo> Estimado Mmanuel, </parrafo>
    <parrafo> Te confirmo nuestra llegada el día 16 en el tren
      de las 8. </parrafo>
    <parrafo> Un saludo </parrafo>
  </contenido>
</fax>
```

©Baltasar Fernández Manóñ

Procesamiento de documentos: distintas presentaciones



Componentes básicas de SGMML

- m Elementos
 - o Información textual con contenido entre marcas
- m Atributos
 - o Información complementaria sobre los elementos
 - Pares nombre-valor dentro de una etiqueta
 - Ej. <precio moneda="dolar"> 100 </precio>
- m Entidades
 - o Modularidad de la información
- m Instrucciones de proceso
 - o Información procedimental que se incluye en el documento marcado
 - o Está destinada a la aplicación que procesa el documento

Utilización de SGML

- m Grandes corporaciones que necesitan gestionar ingentes cantidades de documentos durante largos periodos
 - o Aviónica
 - o Defensa
 - o Editoriales
- m Se han definido DTD estándares para simplificar la compartición de información
 - **Editoriales**: ISO12083 para publicación de libros y revistas
 - **Software**: DocBook para documentación técnica
- m Alto coste y dificultad de implementación
- m Mucha experiencia y probada utilidad

©Baltasar Fernández Manóvil

Estándares relacionados con SGML

- m Lenguaje de especificación y semántica de estilo de los documentos (DSSSL)
 - o Creación de hojas de estilo
 - o Permite especificar transformaciones de documentos
- m Lenguaje de descripción de aplicaciones hipermedia (HyTime)
 - o Mejora la capacidad de representar, direccionar y enlazar información estática y dinámica en cualquier formato o medio
- m Lenguajes de propósito específico
 - o Lenguaje de marcado de vectores (VML)
 - o Lenguaje de marcado para química (CML)

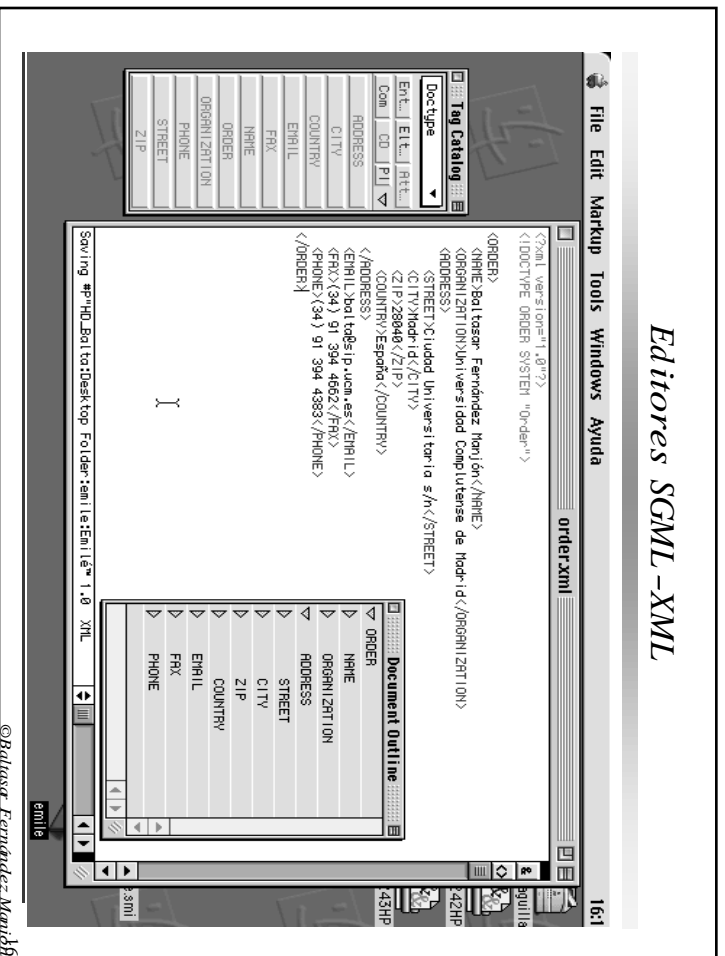
©Baltasar Fernández Manóvil

Lenguaje de marcado extensible (XML)

- m Lenguaje de marcado extensible (XML)
 - o Subconjunto de SGML optimizado para la red
 - o Desarrollo más simple de aplicaciones
- m Hojas de estilo en cascada (CSS1 y CSS2)
 - o Información de estilo para documentos HTML o XML
- m Lenguaje de estilo extensible (XSL)
 - o Basado en DSSSL y CSS
- m Lenguaje de enlace extensible (XLL)
 - o Mayor flexibilidad en la creación de enlaces
- m Lenguaje de consulta de XML (XML-QL)

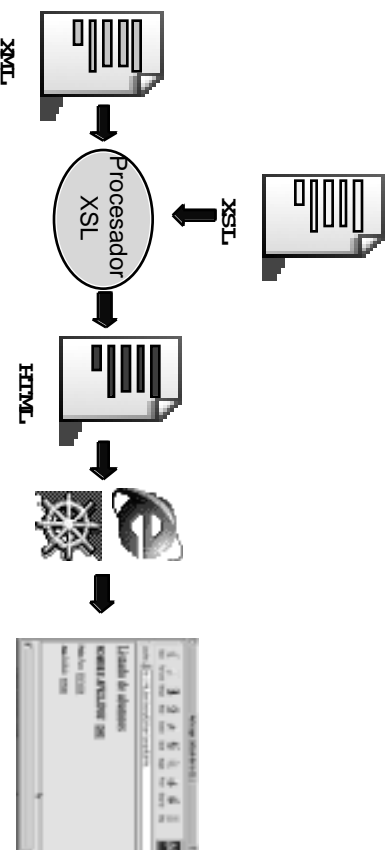
©Balasar Fernández Manó5

Editores SGML/XML



©Balasar Fernández Manó5

Procesamiento documento XML para obtener HTML



©Balasra - Escandalar Manizá7

Marcado e interacción hombre-computadora

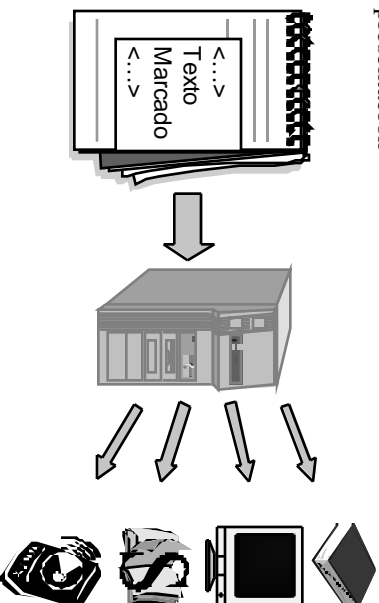
- m Accesibilidad de la información
- m Personalización y adaptación de contenidos
- m Búsqueda “inteligente” de información
- m Extracción de información
- m Comercio electrónico
- m Desarrollo de software

©Balasra - Escandalar Manizá8

Accesibilidad de la información

La separación entre contenido, estructura y presentación permite distintas presentaciones para usuarios discapacitados

- o ICADD (International Committee for Accessible Document Design)
- o Con un único contenido sólo hay que modificar los procesos de presentación



©Baltasar Fernández Manóvil

Accesibilidad y configuración de documentos

Documentos interactivos y configurables por el usuario

Generación automática de ayudas navegacionales a partir de la estructura jerárquica de los documentos

- Tablas de contenidos navegables
- Mapas de contenidos

Distintas presentaciones de un mismo documento

- o Distintas “vistas”
 - Documento con anotaciones
 - u Tres vistas: Sólo documento, sólo anotaciones, ambos contenidos
 - Presentaciones parciales
 - u Documento multilingüe
 - u Sólo la parte del idioma seleccionado
 - Documento multiproducto
 - u Sólo las partes comunes y las específicas de ese producto

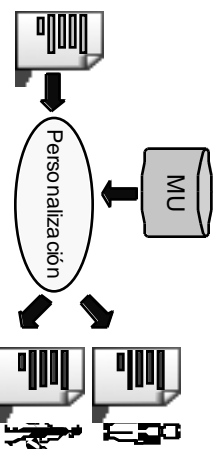
©Baltasar Fernández Manóvil

Personalización de contenidos

- Personalización automática de contenidos y/o interacción a cada usuario concreto
 - Adaptación a las características del usuario
 - Intereses
 - Nivel de conocimiento
 - Preferencias

Aplicación de técnicas de modelado de usuario (MU)

- El sistema adquiere y mantiene la información que necesita sobre cada usuario para realizar la adaptación



©Baltasar Fernández Manóh

Búsqueda “inteligente” de información

Se dispone de grandes cantidades de información en formato electrónico

- Intranets y Web
- Búsqueda de información
 - Mediante palabras claves dentro de todo el documento
 - Navegación a través de los resultados (*Browsing*)
- Problemas
 - Se recuperan muchos documentos que hay que revisar manualmente
 - Muchos de los documentos recuperados no son relevantes

Con información marcada las búsquedas pueden ser más precisas ya que se pueden hacer sólo dentro de las partes significativas de un documento

- Se puede especificar el contexto de la búsqueda
 - Ejemplo: ¿artículos escritos por Bill Gates?
 - Apariciones sólo dentro de elementos marcados como autor

©Baltasar Fernández Manóh

Búsqueda, extracción y filtrado de información

m El marcado permite integrar datos y texto en los documentos sin perder información

m Actualmente

- o Agentes de compra (*shopbots*)

- *BarGainFinder*: Encuentra los precios de CD en diversas tiendas Web

- u Búsqueda muy simple

- o Agentes activos de búsqueda de información

m La disposición de información marcada permitirá la mejora y sofisticación de estas técnicas

- Ej. Búsqueda comparativa de productos en distintos países

- u <precio moneda="dolar">100 </precio >

- u < precio moneda="peseta">15000 </precio >

©Balasar Fernández Manóvil

Comercio electrónico

m Mayor disponibilidad de servicios y mayor simplicidad de uso

m El marcado facilita el intercambio electrónico de datos

- o Se puede expresar información compleja

- o Es fácil realizar conversiones entre distintos formatos

- o Permite integrar bases de datos y aplicaciones Web

m Se puede separar los datos de sus instrucciones de procesamiento

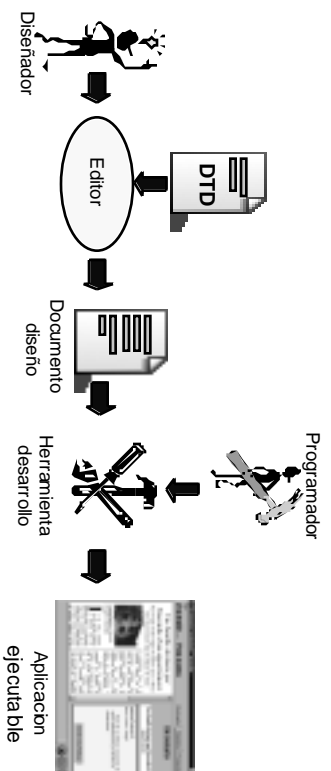
m Potenciado por lenguajes de programación multiplataforma

- o Java

©Balasar Fernández Manóvil

Ayuda al desarrollo de software altamente interactivo

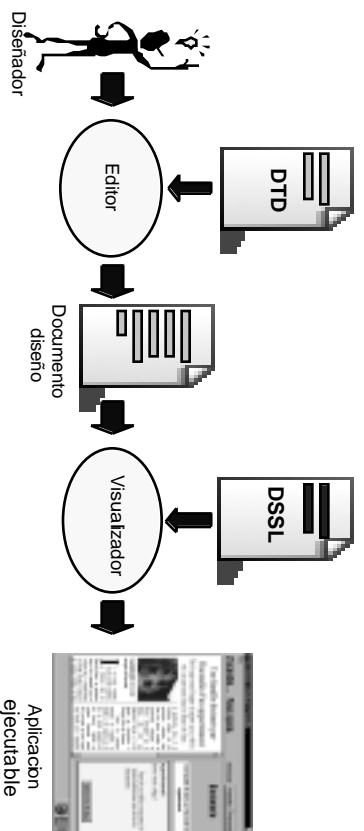
- o Diseño
 - Comunicación diseñadores- programadores
- o Mantenimiento
 - Prueba, depuración, actualización
- o Compatibilidad y dise minación
 - PC, mac, WWW



©Baltasar Fernández Manóñ5

Desarrollo de software altamente interactivo

m ¿Creación automatizada de aplicaciones a partir del diseño?



©Baltasar Fernández Manóñ5

Referencias

- m Campo muy activo y de rápida evolución
 - o Exceso de información
 - o Gran cantidad de estándares, recomendaciones, etc
- m World Wide Web
 - o <http://www.w3.org/>
- m SGML
 - o <http://www.ca-sis-open.org/>
- m XML
 - o <http://xml.org/>

</final>

Muchas gracias por su atención

?Alguna pregunta?

</final>