

Semillero de Introducción a las Tecnologías XML

Universidad de Medellín
Septiembre de 2011

Presentación

- Quienes somos y que queremos aprender
- El Profesor:

William David Velásquez Ramírez

william@bitlab.com.co

<http://bitlab.com.co>

Acercas del Curso

- Su objetivo es que los estudiantes estén en capacidad de aplicar las tecnologías básicas de XML a problemas reales en el ejercicio de su profesión
- Dirigido quienes les guste programar y a quienes no. Los conocimientos previos de HTML van a ayudarles mucho, pero no son indispensables.
- Es un curso práctico, no teórico
- Los estudiantes deben consultar por sí mismos las definiciones formales para todo lo que vamos a ver en la práctica
- Está unido a un segundo semillero de Programación con XML

Introducción

- Por qué y para qué?
 - Porque definitivamente siempre existirán sistemas heterogéneos
 - Porque necesitamos intercambiar información
 - Porque necesitamos distinguir entre información y contenido
- La vida antes de XML
 - Intercambiar información entre sistemas solo podía hacerse a escala local
 - Existe un estándar llamado SGML pero demasiado abstracto

Qué es XML

- El nombre XML viene de las siglas eXtensible Markup Language (Lenguaje eXtensible de Etiquetas).
- XML es un lenguaje para representar información estructurada en forma de documentos. XML permite guardar información en un documento, acompañada de etiquetas que describen la información almacenada en el documento.
- XML se ocupa de la sintaxis para representar la información, pero deja la semántica a criterio del usuario, aunque hoy por hoy existen aplicaciones de XML para describir la semántica de un documento (DTD, XML Schema, OWL, Relax NG...).
- En el mundo real, no se usa simplemente XML, sino alguna de sus aplicaciones.

El origen de XML

- En la década de los 70 IBM desarrolla el GML (General Markup Language) para representar documentos para su sistema de autoedición
- En 1986 ISO acoge el GML y lo convierte en un standard (ISO 8879) y lo renombra como SGML (Standard Generalized Markup Language)
- En 1989 nace HTML siguiendo las normas de SGML. Pronto demuestra problemas para separar el contenido de la presentación y el medio de representación.
- En 1998 Jon Bosak de Sun Microsystems lidera un comité para la creación de XML

Los objetivos de Diseño de XML

- XML debe ser utilizable directamente sobre Internet.
- XML debe soportar una amplia variedad de aplicaciones.
- XML debe ser compatible con SGML.
- Debe ser fácil escribir programas que procesen documentos XML.
- El número de características opcionales en XML debe ser mantenido en un mínimo, idealmente cero.
- Los documentos XML deben ser legibles por un humano y razonablemente claros.
- El diseño de XML debe ser preparado rápidamente

Objetivos de Diseño (2)

- El diseño de XML debe ser formal y conciso.
- Los documentos XML deben ser fáciles de crear.
- La brevedad en la marcación es de mínima importancia
- Esta especificación, junto con los estándares asociados (Unicode and ISO/IEC 10646 para caracteres, Internet RFC 1766 para las marcas de identificación de lenguaje, ISO 639 para los códigos de nombre de lenguaje, ISO 3166 para los códigos de nombre de país), provee toda la información necesaria para entender XML Version 1.0 y construir programas de computador que lo procesen.

Las ventajas de XML

- Redundancia
- Auto descriptivo
- Capacidad de representar la estructura humana de la información
- Diseñado para la red

Algunos usos actuales de XML

- Páginas Web: XHTML y XSLT
- Suscripciones y Podcasts: RSS y Atom
- Interfaces de Programación: SOAP y REST
- Aplicaciones de Oficina: ODF y OpenXML
- Pago de Impuestos: DIAN
- Negocios Electrónicos: ebXML, UBL, XBRL

**Documentos
Electrónicos**

**XSLT-FO
DocBook
DITA
XPS
OOXML
ODF**

Web

**XHTML
RSS
Atom
SVG**

**Negocios
Electrónicos**

**ebXML
UBL
EDI-XML
XBRL**

**Aplicaciones
Distribuidas**

**SOAP
WSDL
UDDI
REST
XML-RPC
BPEL**

**Usos
Particulares**

Estructura: DTD, XML Schema, RelaxNG, Schematron

Referencia: Xpath, XLink, XPointer

Transformación: XSLT, XSLT, XQuery

Seguridad: XML Signatures

Conceptos

```
<?xml version="1.0" encoding="UTF-8" standalone="yes" ?>
<!DOCTYPE hola SYSTEM "hola.dtd" ?>
<?xslt stylesheet="xslt.xsl" ?>
<!-- Este es un ejemplo de un documento XML -->
<hola>
  <nombre nombre="Tiene" ?>Saludos Damas & Caballeros!</nombre>
  <![CDATA[Aqui < > allá]]>
</hola>
```

The diagram illustrates various XML concepts with callouts pointing to specific parts of the code:

- XML Declaration**: Points to the `<?xml version="1.0" encoding="UTF-8" standalone="yes" ?>` line.
- DOCTYPE Section**: Points to the `<!DOCTYPE hola SYSTEM "hola.dtd" ?>` line.
- Processing Instruction**: Points to the `<?xslt stylesheet="xslt.xsl" ?>` line.
- Root Node**: Points to the `<hola>` opening tag.
- Atributos**: Points to the `nombre="Tiene"` attribute in the `<nombre>` tag.
- Elementos**: Points to the `<nombre>` tag.
- CDATA Section**: Points to the `<![CDATA[Aqui < > allá]]>` section.
- Entity**: Points to the `&` character in the text `Saludos Damas & Caballeros!`.
- Text**: Points to the text `Saludos Damas & Caballeros!`.

Notación XML

- Documento
- Etiquetas (Tags)
- Start Tag
- End Tag
- Empty Tag

Notación XML (2)

- Elementos
- Elemento Raíz
- Atributos
- Processing Instructions
- XML Declaration
- Encoding
- CDATA
- Comentarios
- DOCTYPE

Los temas de XML

- Encoding
- Representación de Información Binaria
- Prefijos y Namespaces

Práctica

- Escriba su hoja de Vida Usando XML
 - Nombres y apellidos
 - Direccion telefono email
 - Fecha de nacimiento
 - Estudios
 - Institucion
 - Curso
 - Año de Finalizacion
 - william@bitlab.com.co

Mejores Practicas

- XML no es HTML
- Mayúsculas != minúsculas
- Use siempre comillas para los atributos
- Cierre todo lo que abra y en el mismo orden
- Al programar, tenga muy en cuenta el encoding

Elementos o Atributos?

- Los elementos pueden contener otros elementos o atributos
- Los elementos pueden repetirse
- Los atributos ocupan menos espacio
- Los atributos son opcionales por defecto
- Ambos pueden tener restrictions, facets o enumerations

Para saber más

- W3C World Wide Web Consortium
<http://www.w3c.org/xml>
- O'Reilly Publications: www.xml.com
- [Www.w3schools.com](http://www.w3schools.com)