

XML

eXtensible Markup Language

Ing. Ramón Roque Hernández, M.C.
ramonroque@yahoo.com
<http://www.geocities.com/ramonroque>

XML

Introducción

¿Qué es XML?

- Es un formato simple, plano, flexible y estándar para describir datos.
- Fue diseñado, desarrollado y es mantenido por el World Wide Web Consortium (W3C).
- Ideado para ser un formato universal.
- A partir de febrero 10 de 1998, XML se convirtió en una recomendación oficial del W3C.

¿Qué es el W3C?

- W3C = World Wide Web Consortium
- Es un organismo internacional cuyo objetivo es:
“Guiar la Web hacia su máximo potencial a través del desarrollo de protocolos y pautas que aseguren el crecimiento futuro de la Web”.
- El W3C Desarrolla estándares WEB.
- Creado en 1994 por Tim-Berners-Lee (Creador del WWW).
<http://www.w3c.es> (Página en español)
<http://www.w3.org> (Página en inglés)

Origen del XML

- El predecesor de XML es el SGML (Standard Generalized Markup Language).
- SGML proporcionaba una manera consistente de aplicar etiquetas para la descripción de secciones en un documento.
- Se publicó en un amplio documento con alto grado de dificultad. No se entendió ni se implementó completamente.
- Tiempo después se eliminaron de este lenguaje los componentes más complejos y menos útiles, manteniendo la misma filosofía, y de esta manera nació XML.

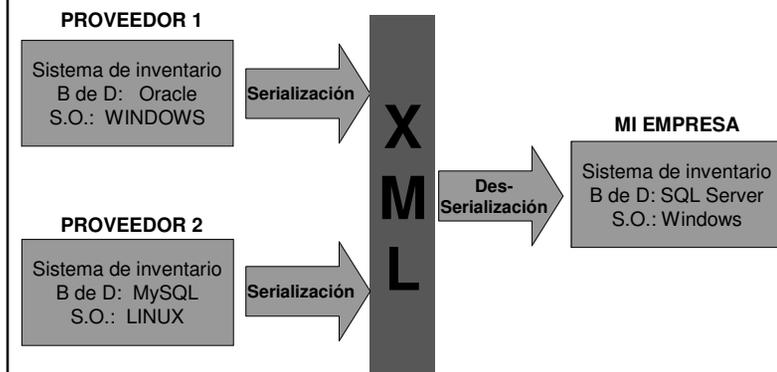
Generalidades de XML

- XML está basado en el uso de etiquetas.
 - Ejemplo:
`<Nombre>Ramon</Nombre>`
- Cada persona define sus etiquetas
- XML permite crear nuevos lenguajes a partir de él.
- El uso de XML permite pensar acerca de la información en contenido, estructura y presentación de manera separada.

XML en el intercambio de datos

- XML permite el intercambio de datos entre diferentes plataformas y formatos, ya que permite la descripción y entrega de datos en una manera estructurada, consistente, y totalmente independiente de hardware y software.
- Una aplicación utiliza la “serialización” para generar XML a partir de objetos internos propios; y la “des-serIALIZACIÓN” para reconstruir internamente objetos provenientes de XML.

XML en el intercambio de datos



XML en el WEB

- **PROBLEMA:**
 - El WEB requiere un formato simple, ligero, que ocupe poco espacio para la transferencia de datos.
 - **SOLUCION:**
 - XML
- Ejemplo:
- **Protocolo SOAP**
 - Simple Object Access Protocol
 - Protocolo de comunicación entre aplicaciones a través de HTTP, basado en XML.
 - Diseñado especialmente para internet

XML vs. EDI

- XML representa una nueva y mejorada solución al antiguo EDI (Electronic Data Interchange).
- XML es mas sencillo y comprensible que EDI.



XML vs. HTML

- XML y HTML son diferentes.
- **XML**
 - Diseñado para describir datos centrándose en su contenido
- **HTML**
 - Diseñado para mostrar datos centrándose en su apariencia.
- Ambos tienen en común:
 - El uso de etiquetas
 - Se almacenan en formato plano (texto).
- XML no fue creado para sustituir a HTML.
- Ambos ofrecen distinto valor agregado a las aplicaciones.

XML Anatomía, Validación y Desplegado

Anatomía de XML

Instrucciones de procesamiento

Comentarios

Elemento Raíz (Contiene todo el documento)

```
<?xml version="1.0" encoding="UTF-8" ?>
<!-- Este es el archivo de Shippers -->
<dataroot company="NorthWind">
  <Shippers>
    <ShipperID>1</ShipperID>
    <CompanyName>Speedy Express</CompanyName>
    <Phone>(503) 555-9831</Phone>
  </Shippers>
  <Shippers>
    <ShipperID>2</ShipperID>
    <CompanyName>United Package</CompanyName>
    <Phone>(503) 555-3199</Phone>
  </Shippers>
</dataroot>
```

Elementos hijo

Documento bien formado

- Es aquel que está escrito adecuadamente (completo) y posee una sintaxis correcta. Ejemplo:

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<catalog>
  <cd>
    <title>Grandes Exitos</title>
    <artist>Artistas Varios</artist>
    <country>Mexico</country>
    <company>EMI</company>
    <price>200.00</price>
    <year>2005</year>
  </cd>
</catalog>
```

Reglas para documentos bien formados

- Todos los elementos deben tener una etiqueta de apertura y una de cierre.
- Los nombres de las etiquetas deben ser congruentes en el uso de mayúsculas y minúsculas; no deben contener espacios en blanco y no deben iniciar con números, signos de puntuación o la palabra "xml".
- Los elementos deben estar anidados correctamente.
- Todo documento debe tener un solo elemento raíz que contenga al resto de los elementos.
- Los espacios en blanco son respetados y conservados.
- Los comentarios se colocan entre <!-- y -->
- Las instrucciones de procesamiento deben aparecer una sola vez al inicio del documento, y comenzar siempre con la palabra "xml".

Parsers

- Un Parser XML (también llamado procesador) es un programa escrito para leer, actualizar, y crear contenido XML.
- Las aplicaciones que generan o leen datos en formato XML requieren de él para cumplir con su función.
- Tareas que realiza un parser:
 - Manipular y verificar que el contenido de un documento XML esté bien formado.
 - Encontrar, extraer y utilizar el contenido de nodos específicos
 - Determinar el tipo de nodos
 - Ordenar nodos por sus espacios de nombre.
 - Algunos parsers permiten validar el documento de acuerdo a un esquema.

Parsers

- Actualmente existen muchos parsers con características variadas, y que se ejecutan en diversas plataformas.
- Muchos son gratuitos.
- Un buen catálogo de parsers XML se encuentra en:
http://www.xml.com/pub/rg/XML_Parsers

Validación

- **Documento Válido**
 - Es un documento bien formado que además se apega a las reglas de construcción definidas en un archivo separado.
- **“Validar”**
 - Es comparar la estructura, nombres y tipos de datos de un documento XML con un conjunto de requerimientos.
- **¿Cuándo validar?**
 - Siempre que se transfiera información entre sistemas u organizaciones
 - Cuando se almacenan datos provenientes de un módulo de captura.

Formas de Validar: DTD y XSD

- **DTD**
 - Document Type Definition
 - Forma antigua de validar un documento XML
- **Esquema (XSD)**
 - XML Schema Definition
 - Nueva forma estándar de validar un documento XML

Validación



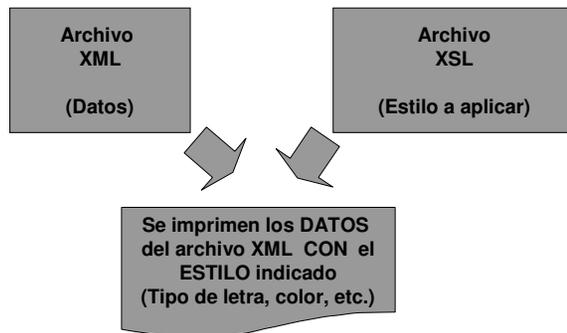
Desplegado sin estilo de archivos XML

- Un documento XML bien formado puede desplegarse sin problemas en el Internet Explorer.
- Se Muestra el archivo de forma jerárquica y permite expandir/contrair sus elementos.
- Si el archivo XML tiene errores, el Internet Explorer indica su ubicación.

Desplegado con estilo de archivos XML

- CSS
 - (**C**ascade **S**tyle **S**heet)
 - Forma antigua de poner estilos
- XSL
 - (**e**Xtensible **S**tylesheet **L**anguage)
 - Nuevo estándar para poner estilos

Desplegado con estilos



The image shows three windows. The top window is "ARCHIVO XML (DATOS)" in Notepad, containing XML code for a CD collection. The middle window is "ARCHIVO XSL (ESTILO)" in Notepad, containing XSL code that defines a table for the CD data. The right window is "DESPLEGADO DE DATOS CON ESTILO" in Microsoft Internet Explorer, showing the rendered output: a table with columns "Titulo" and "Artista" and a row "Grandes Exitos | Artistas Varios".

XML **Importancia** **en la informática**

Importancia actual del XML

- SQL Server y Oracle han agregado características avanzadas en sus mas recientes versiones para la representación, conversión y manejo nativo de datos XML.
- La siguiente versión de MICROSOFT OFFICE almacenará todos sus archivos en XML.
- INTEL, CISCO, DATAPOWER incluyen ya XML en sus aplicaciones.
- XML es la base para el desarrollo de servicios WEB en VISUAL STUDIO.NET 2005 utilizando SOAP.
- Muchas tecnologías de actualidad se relacionan con XML.

Algunas tecnologías **relacionadas con XML**

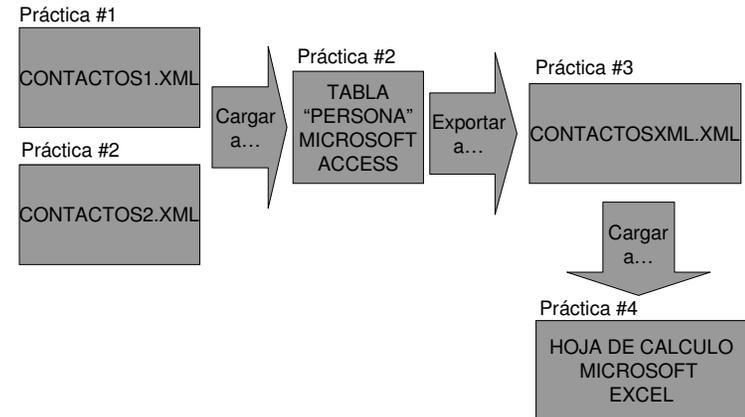
- **XPATH**
 - Lenguaje para encontrar información en un documento XML.
- **XQUERY**
 - Lenguaje de consultas simples y complejas parecido al SQL para extraer datos de un documento XML.
- **XHTML**
 - Versión estricta y mejorada de HTML
- **XFORMS**
 - Nueva generación de controles de captura de datos en las formas WEB.

Algunas tecnologías **relacionadas con XML**

- **DOM**
 - (Document Object Model)
 - Modelo para representar y manipular documentos XML desde un programa cualquiera.
- **RSS**
 - (Really Simple Syndication)
 - Estándar para compartir encabezados de noticias y contenido.
- **WAP**
 - (Wireless Application Protocol)
 - Protocolo de comunicación para dispositivos móviles.

XML Prácticas

Vista Preliminar de lo que se hará en las prácticas

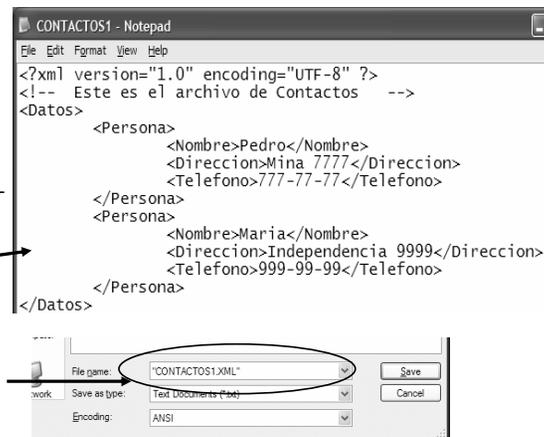


Práctica #1 Creando XML desde un editor de textos

1. Entrar al BLOQUE DE NOTAS

2. Teclear el archivo XML (Ser cuidadoso con mayúsculas y minúsculas)

3. Guardarlo como "CONTACTOS1.XML"



4. Ubicar el archivo en el Explorador de Windows, y dar doble CLICK.



Si existe algun error, el Internet Explorer lo indicará y se deberá corregir el archivo desde el block de notas.

5. Se abre el Internet Explorer. Observar como se pueden expandir/contrar los elementos del archivo.



Práctica #2

Cargando a una base de datos el contenido de dos archivos XML

Se usará el archivo XML de la práctica #1.

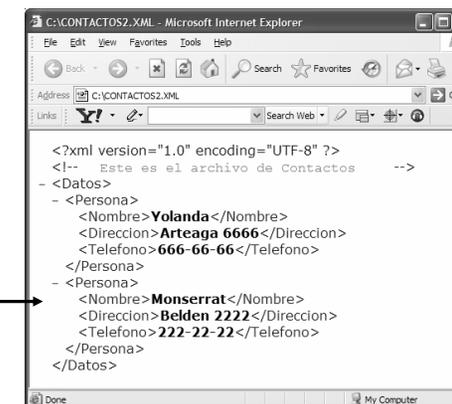
1. Además, se debe crear otro archivo en el mismo directorio que el anterior.
2. Entrar al Block de Notas
3. Guardarlo como "CONTACTOS2.XML"

4. Ubicar el archivo en el Explorador de Windows, dar doble click.



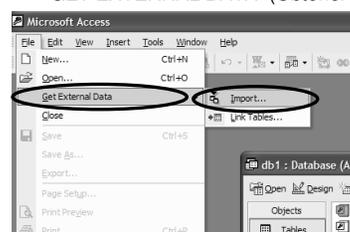
Si existe algún error, el Internet Explorer lo indicará y se deberá corregir el archivo desde el block de notas.

Se abre el Internet Explorer. Observar como se pueden expandir/contrair los elementos del archivo.



5. Entrar a MICROSOFT ACCESS. Crear una nueva Base de Datos en BLANCO.

6. En el menú FILE (Archivo), GET EXTERNAL DATA (Obtener datos externos), IMPORT (Importar)...



7. Seleccionar "CONTACTOS1.XML"



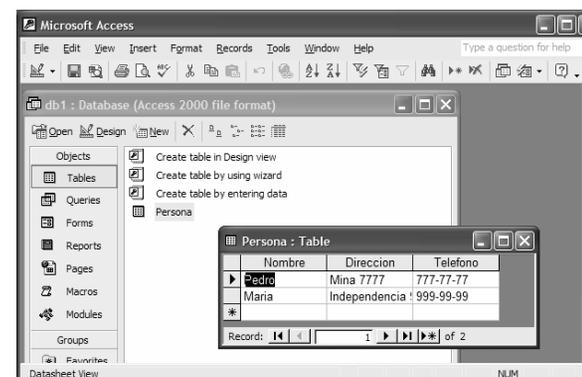
8. Aparece la siguiente pantalla, seleccionar el botón OPTIONS (Opciones) Marcar STRUCTURE AND DATA (Estructura y datos).



9. Presionar OK (Aceptar). Aparece una pantalla informando Que el proceso terminó:



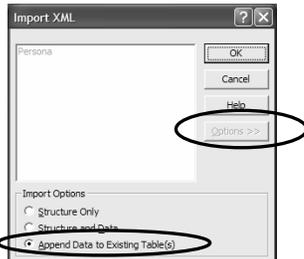
10. ACCESS crea una tabla con la estructura y los datos del archivo XML



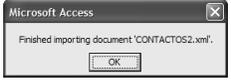
11. En la **misma base de datos**, Repetir el proceso con el segundo archivo:
 En el menú FILE (Archivo),
 GET EXTERNAL DATA (Obtener datos externos),
 IMPORT (Importar)... Seleccionar CONTACTOS2.XML



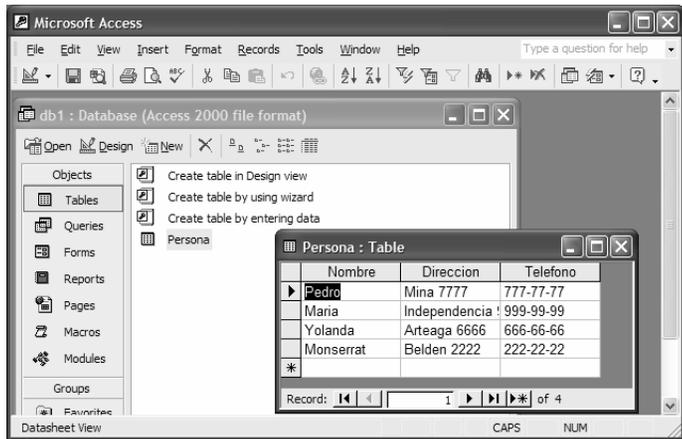
12. Aparece la siguiente pantalla, seleccionar el botón OPCIONES....
 Marcar "APPEND DATA TO EXISTING TABLES"
 (Agregar datos a tablas existentes)



13. Presionar OK (Aceptar).
 aparece una pantalla informando que el proceso terminó:



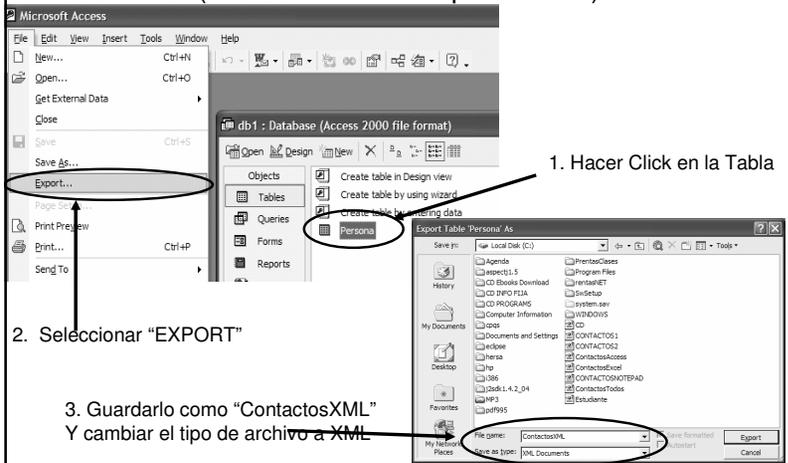
14. Los datos de "CONTACTOS2.XML" fueron importados y agregados al contenido De la tabla "Persona" que ya contenía los datos de "CONTACTOS1.XML"



Nombre	Direccion	Telefono
Pedro	Mina 7777	777-77-77
Maria	Independencia	999-99-99
Yolanda	Arteaga 6666	666-66-66
Monserrat	Belden 2222	222-22-22

Práctica #3

Exportando de Access a XML (Continuación de la práctica #2)

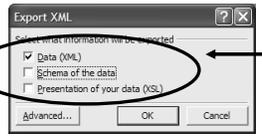


1. Hacer Click en la Tabla

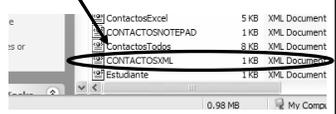
2. Seleccionar "EXPORT"

3. Guardarlo como "ContactosXML"
 Y cambiar el tipo de archivo a XML

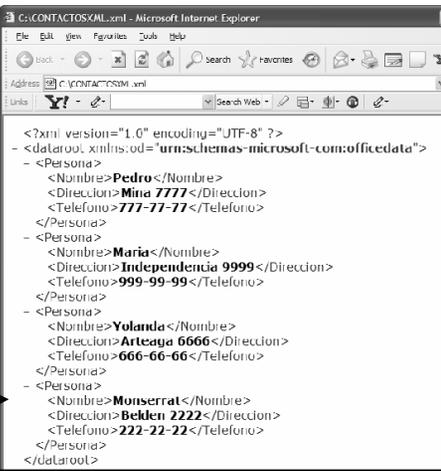
4. Seleccionar solo la primera opción
 Presionar OK (Aceptar)



5. En este momento, se ha creado ya un archivo llamado "CONTACTOSXML.XML" en el directorio indicado.
 Con el Explorador de Windows ubicarlo y dar doble click.



6. Debe aparecer el Internet Explorer desplegando el archivo con todos los datos.

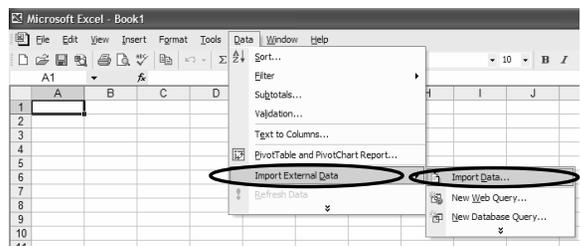


```

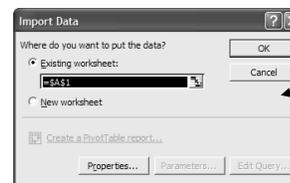
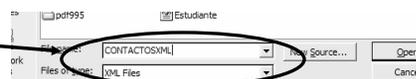
<?xml version="1.0" encoding="UTF-8" ?>
<dataroot xmlns:od="urn:schemas-microsoft-com:officedata">
  <Persona>
    <Nombre>Pedro</Nombre>
    <Direccion>Mina 7777</Direccion>
    <Telefono>777-77-77</Telefono>
  </Persona>
  <Persona>
    <Nombre>Maria</Nombre>
    <Direccion>Independencia 9999</Direccion>
    <Telefono>999-99-99</Telefono>
  </Persona>
  <Persona>
    <Nombre>Yolanda</Nombre>
    <Direccion>Arteaga 6666</Direccion>
    <Telefono>666-66-66</Telefono>
  </Persona>
  <Persona>
    <Nombre>Monserrat</Nombre>
    <Direccion>Belden 2222</Direccion>
    <Telefono>222-22-22</Telefono>
  </Persona>
</dataroot>
    
```

Práctica #4 Importando XML a EXCEL

1. Entrar a EXCEL
2. Seleccionar DATA, "Import External Data", "Import Data"

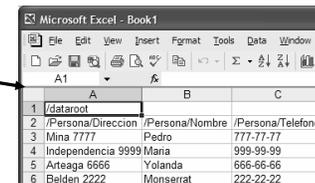


3. Ubicar y seleccionar "CONTACTOSXML.XML" Creado en la practica #3



4. Seleccionar el lugar donde se desean Poner los datos en la hoja de EXCEL

5. EXCEL carga a la hoja los datos del archivo XML



XML Conclusiones

Conclusiones

- XML es:
 - Un estándar universal, basado en etiquetas.
 - Almacenado en formato de texto plano.
 - Es independiente de hardware y software.
- Permite a las aplicaciones superar las limitaciones de compatibilidad de las plataformas en las que fueron desarrolladas.
- Importante principalmente en dos escenarios:
 - Intercambio de datos entre aplicaciones.
 - WEB.
- XML separa el contenido, los esquemas y el despliegado de los datos en archivos separados.
- Las tecnologías relacionadas con XML son las que exponen su verdadera importancia.
- XML será tema cotidiano por los próximos años.