

# Data and Information Management through XQUERY, XML Y XSL Technologies

## Manejo de los Datos y la Información Mediante las Tecnologías XQUERY, XML Y XSL

J C. Calabria, A Pico, F Thomas

**Keywords:**

Data, DTD, Information system, Technologies, XML, XSL, XQUERY.

**Abstract**

In this document there is described the use of the technologies of the information in particular case of those that work with the information and the information. The metalanguage XML is turning into the base for the construction of the great majority of the applications and documents, the XSL is A XSL's part his intention is, to organize the structure of a document XML doing that the persons and you scheme could understand them easily. These technologies are born to unify the form in which the applications handle the information and the information so much own, as his communication with other applications be or not similar. Since at present the most relevant actors in the organizations are the information and information for this way this type of technologies will be those that will order the stop in the systems of information of the current and new organizations, being like that most used by his potential, processing, agility between many other characteristics. It is necessary to clarify that these technologies in day are a very important advance for the management of the information and his presentation with a limited cost and with the advantages of achieving results in short term. Of equal form, the article describes the importance of the managing of the information for the managers of information, archivists and engineers.

**Palabras clave:**

Datos, DTD, Sistema de Información, Tecnologías, XML, XSL, XQUERY.

En este documento se describe el uso de las tecnologías de la información en caso particular de aquellas que trabajan con la información y los datos. El metalenguaje XML se está convirtiendo en la base para la construcción de la gran mayoría de las aplicaciones y documentos, el XSL es parte de XSLT su propósito es, organizar la estructura de un documento XML haciendo que las personas y maquinas puedan entenderlos fácilmente. Estas tecnologías nacen para unificar la forma en que las aplicaciones manejan los datos y la información tanto propios, como su comunicación con otras aplicaciones sean o no similares.

Ya que en la actualidad los actores más relevantes en las organizaciones son

los datos e información por este modo este tipo de tecnologías serán las que mandarían la parada en los sistemas de datos de las actuales y nuevas organizaciones, siendo así las más usadas por su potencialidad, procesamiento, agilidad entre muchas otras características. Cabe aclarar que estas tecnologías en día son un avance muy importante para la gestión de los datos y su presentación con un costo reducido y con las ventajas de lograr resultados en corto plazo. De igual forma, el artículo describe la importancia del manejo de la información para los gestores de información, archivistas e ingenieros.

### I. INTRODUCCION

Este artículo trata de dar a conocer tres de las tecnologías para el manejo de la información, las cuales son XML, XSL y XQUERY. El metalenguaje XML o lenguaje que describe otros lenguajes, en el que se diseñan marcas con el fin de dar estructura a la información y de esta manera poder construir una gran variedad de documentos que sean fácilmente entendidos por personas y maquinas. El papel que desempeña el XSL en el manejo de la información es el de transformar documentos XML donde el documento original no se modifica, por lo que un nuevo documento se crea basándose en el contenido de uno ya existente, El nuevo documento puede ser la salida del transformador en la sintaxis estándar de XML, como HTML o texto plano. En la actualidad XSL es la más utilizada para convertir los datos entre los diferentes esquemas XML o para convertir los datos XML en páginas web o PDF de los documentos. Por su parte XQuery es un lenguaje de consulta para documentos XML, con funcionalidades similares a SQL. Lo más simple es pensar que XQuery es a XML como SQL es a las tablas relacionales.

El metalenguaje XML se está convirtiendo en la base para la construcción de la gran mayoría de las aplicaciones y documentos para que estos sean fácilmente entendidos por las personas o maquinas, convirtiéndose XML en una norma para realizar una gestión eficiente de la información local y distribuida. Esta tecnología también ha nacido para unificar la forma en que las aplicaciones manejan los datos e información, tanto propios como su comunicación con otras aplicaciones sean o no similares.

### II. DESARROLLO DEL ARTÍCULO

En esta sección describiremos todos los conceptos manejados en los sistemas de información y datos.

**II-A Sistema De Información (SI):** es un conjunto de elementos orientados al tratamiento y administración de datos e información, organizados y listos para su posterior uso, generados para cubrir una necesidad. Dichos elementos formarán parte de alguna de estas categorías:

Elementos de un sistema de información:

- Personas.
- Datos.
- Actividades o técnicas de trabajo.
- Recursos materiales en general (típicamente recursos informáticos y de comunicación, aunque no tienen por qué ser de éste tipo obligatoriamente).

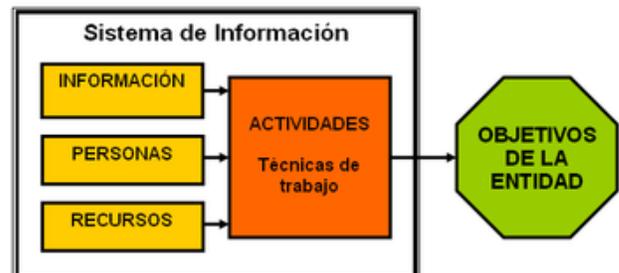


Fig. 1 Sistema de información

**II-B XML:** (Extensible Markup language). Es un metalenguaje extensible de etiquetas desarrollado por el World Wide Web Consortium (W3C). Es una simplificación y adaptación del SGML y permite definir la gramática de lenguajes específicos (de la misma manera que HTML es a su vez un lenguaje definido por SGML). Por lo tanto XML no es realmente un lenguaje en particular, sino una manera de definir lenguajes para diferentes necesidades. Algunos de estos lenguajes que usan XML para su definición son XHTML, SVG, MathML, Android.

**II-C DTD:** (Document Type Definition). Es una descripción de estructura y sintaxis de un documento XML o SGML. Su función básica es la descripción del

formato de datos, para usar un formato común y mantener la consistencia entre todos los documentos que utilicen la misma DTD. De esta forma, dichos documentos, pueden ser validados, conocen la estructura de los elementos y la descripción de los datos que trae consigo cada documento, y pueden además compartir la misma descripción y forma de validación dentro de un grupo de trabajo que usa el mismo tipo de información.

**II-D XSL:** Es un estándar de la organización W3C que presenta una forma de transformar documentos XML en otros e incluso a formatos que no son XML. Las hojas de estilo XSL. Aunque el término de hojas de estilo no se aplica sobre la función directa del XSL -realizan la transformación del documento utilizando una o varias reglas de plantilla. Estas reglas de plantilla unidas al documento fuente a transformar alimentan un procesador de XSL, el que realiza las transformaciones deseadas poniendo el resultado en un archivo de salida, o, como en el caso de una página web, las hace directamente en un dispositivo de presentación tal como el monitor del usuario. Actualmente, XSL es muy usado en la edición web, generando páginas HTML o XHTML. La unión de XML y XSL permite separar contenido y presentación, aumentando así la productividad.

**II-E XQUERY:** Es un lenguaje de consulta diseñado para colecciones de datos XML. Es semánticamente similar a SQL, aunque incluye algunas capacidades de programación. XQuery 1.0 fue desarrollado por el grupo de trabajo de Consulta XML del W3C. El trabajo fue estrechamente coordinado con el desarrollo de XSL 2.0 por el Grupo de trabajo XSL. Los dos grupos compartieron la responsabilidad del XPath 2.0 que es un subconjunto de XQuery 1.0. XQuery 1.0 es una Recomendación del W3C desde el 23 de enero de 2007.

### III. CARACTERÍSTICAS DE LAS TECNOLOGÍAS DEL MANEJO DE DATOS E INFORMACIÓN

Entre las características que poseen las tecnologías que manejan los datos y la información podemos señalar que mientras el HTML es un lenguaje diseñado con el propósito de mostrar documentos a las personas, ejemplo de estas son páginas web con estilos agradables, por otro lado XML está encaminado al manejo de la estructura formal que componen los

documentos, los XSL por su parte muestran el contenido del XML y los XQUERY permiten realizar las consultas en los XML.

Entre las características más importantes de cada tecnología descrita en la sección siguiente, tenemos las siguientes:

#### A. XML

- Independencia de los datos respecto a las aplicaciones: XML es independiente de cualquier código patentado y formato propietario. XML es libre de usar, sin restricciones ni licenciamientos, con XML se ofrece una gestión de la información totalmente abierta, ya que facilita un formato común normalizado para todas las aplicaciones, de fácil lectura e interpretación tanto para las personas como para las máquinas, adicionalmente XML permite codificar la información utilizando el estándar internacional UNICODE (UTF-8, ISO-8859-1, etc.).
- Información en un contexto: XML fue creado especialmente para que las máquinas puedan entender la semántica y la estructura de un documento y así poder procesarla según las instrucciones que se brinden en el mismo documento o que el sistema sea capaz de realizar.
- Elementos describen la estructura del documento: con XML se puede describir totalmente la estructura de un documento, puede modelar cualquier estructura de acuerdo a las necesidades particulares de las organizaciones y las tipologías documentales asociadas, incluso a partir de XML, se pueden modelar y estructurar todos los documentos que se trabajan en un cuadro de clasificación documental –CCD. Cuando se construye un documento de XML, se requiere, que este sea de fácil entendimiento y lectura tanto para las personas como para las máquinas, además su estructura debe acomodarse de una forma lógica para representar la información lo más organizadamente posible.
- Permite organizar la información siguiendo una estructura ordenada.
- Permite garantizar la validez del documento.

- Posee métodos para visualizar la información: XML contempla el uso de varios lenguajes para representar y transformar la información de acuerdo al contexto que se necesita, tanto para los mismos humanos como para las maquinas, también se pueden visualizar la información contenida ya sea un documento completo o alguna de sus partes.

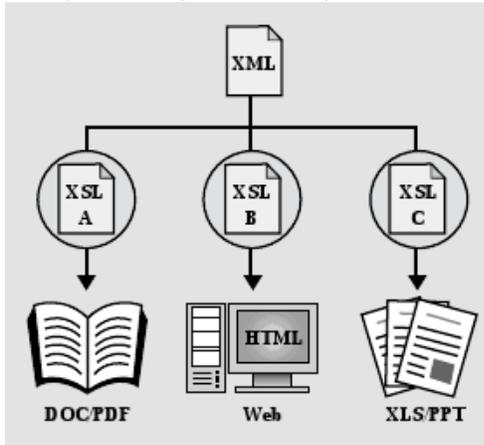


Fig. 2. Transformación De Un Documento A XML

- Posee campos de información y control para las aplicaciones: dentro de XML se pueden colocar componentes de información exclusivos para el uso de las aplicaciones, esto incluye la codificación de los caracteres, la versión del documento, fechas, restricciones de uso, transformaciones del documento, entre otros.
- Mezclar elementos de diferentes lenguajes. Es decir que los lenguajes sean extensibles.
- Permite enlazar datos relacionados: XML fue diseñado sobre la base del hiper-documento sobre el cual los documentos no manejan estructura lineales, como lo manejan HTML, cuando puede saltar en cualquier parte de un documento a otro documento o alguna otra parte del mismo.

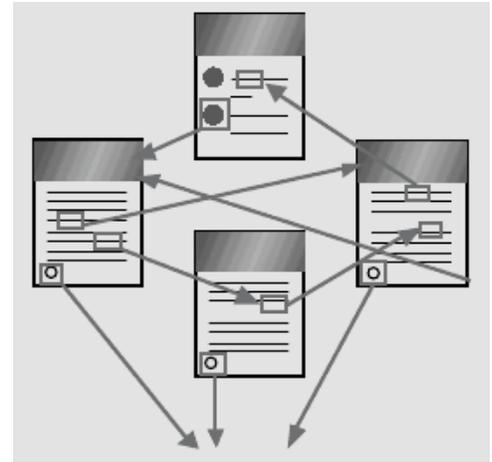


Fig. 3. Hiperdocumento XML

- Ofrece la libertad de formato y medio: XML maneja un lenguaje neutro con posibilidad total de transformación, esto permite convertir un documento XML en otro de forma casi transparente, a su vez permite reutilizar los componentes y documentos que ya existen para generar nuevos documentos que se ajusten a cualquier necesidad de información.
  - La creación de analizadores simples, sin ninguna lógica especial para cada lenguaje.
  - Empezar de cero y hacer hincapié en que no se acepte nunca un documento con errores de sintaxis.
- B. XQUERY*
- XQuery proporciona los medios para extraer y manipular información de documentos XML, o de cualquier fuente de datos que pueda ser representada mediante XML como, por ejemplo, bases de datos relacionales o documentos ofimáticos.
- C. XSL*
- Formatear los elementos fuente basados en relaciones de ancestro/descendiente, posición y unicidad.
  - La creación de construcciones de formato sofisticadas incluyendo texto generado e imágenes
  - La definición de macros de formateo reutilizables
  - Estilos independientes de la dirección en que se escriba el lenguaje
  - Conjunto de objetos de formato extensible

- Una manera de describir el proceso de transformación es decir que las XSL transforma un árbol XML de entrada e otro árbol XML de salida.

IV. APLICACIONES DE LAS TECNOLOGÍAS DEL MANEJO DE DATOS E INFORMACIÓN

Las aplicaciones que tienen las tecnologías que manejan los datos y la información como XSL, XML y XQUERY son muchas y se dan en diferentes áreas de desempeño. A continuación mencionaremos las más importantes:

A. XSL

- Las aplicaciones suelen usar XSL para convertir datos XML en documentos HTML o XHTML para la visualización de una página web. La transformación puede ocurrir de forma dinámica, ya sea en un cliente o en un servidor, o puede ser hecho como parte del proceso de publicación. También se utiliza para crear una salida para la impresión directa o pantalla de vídeo, por lo general mediante la transformación del XML original en objetos de formato XSL para crear formato de salida que luego se puede convertir en una gran variedad de formatos, incluyendo PDF, PostScript y PNG. XSL también puede traducir los mensajes XML entre los diferentes esquemas XML, o hacer cambios a los documentos en el ámbito de un solo esquema, por ejemplo, eliminar los restos de un mensaje que no son necesarios.

B. XQUERY

- Extraer información de una base de datos para usarla en un Servicio Web.
- Generar un resumen de la información almacenada en una base de datos XML.
- Realizar búsquedas textuales en la web y compilar los resultados de la misma.
- Seleccionar y transformar datos de XML a XHTML de forma que se puedan publicar en la Web.
- Obtener datos desde diferentes fuentes con vistas a ser integradas por la aplicación.
- Dividir un documento XML que representa una serie de múltiples transacciones en varios documentos XML, uno por cada transacción.

C. XML

La motivación original que impulso la aparición de XML fue la necesidad de la creciente globalización en los negocios electrónicos, las cuales necesitan por definición del intercambio de datos. De esta forma, XML se ha convertido en la base para la construcción de aplicaciones, brindando la posibilidad de manejar contenidos en documentos estructurados y distribuyéndolos e inter-operándolos en redes de computadores.

Las aplicaciones de XML se pueden focalizar en dos grandes grupos: sistemas de gestión documental y sistemas transaccionales. La aplicación de XML en la gestión documental relacionada principalmente con los sistemas de información que manejan las bibliotecas y archivos, lugares donde se concentran gran cantidad de documentos cuya diseminación al usuario es parte vital del proceso, no importando donde se encuentre este. La gestión documental también puede equipararse a un planteamiento más amplio enfocado a la gestión global de los recursos de información de una institución o empresa mediante tecnologías Web, de esta forma el uso de XML en la gestión documental involucra procesos de creación, estructuración, edición, publicación entre otros.

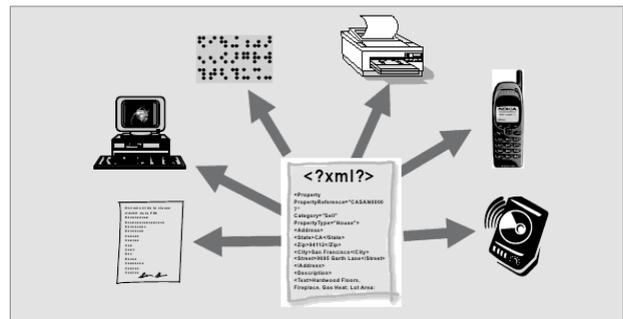


Fig. 4. Usos del XML

Por otra parte en los sistemas transaccionales, las aplicaciones son muchas entre ellas tenemos, transacciones electrónicas, soporte a clientes, personalización de la web, independencia del dispositivo de acceso a la información, gestión de la información/conocimiento, buscador web, intercambio de información

EJEMPLO DEL USO DE LAS TECNOLOGÍAS DEL MANEJO DE DATOS E INFORMACIÓN

El siguiente ejemplo consta de las ventas en un café internet, el cual ofrece servicios de llamadas,

impresión, internet, scanner. Se desea conocer cuáles fueron las ventas al final del día.

El XML correspondiente al ejercicio es el siguiente:

```
<detalle_servicios>
  <dia>
    <valor>2008/03/25</valor>
    <servicios>
      <servicio>
        <tipo_servicio>internet</tipo_servicio>
        <descripcion>
          nro_pc = PC2
          cant_min = 15
        </descripcion>
        <valor>500</valor>
      </servicio>
      <servicio>
        <tipo_servicio>internet</tipo_servicio>
        <descripcion>
          nro_pc = PC1
          cant_horas = 2
        </descripcion>
        <valor>3000</valor>
      </servicio>
    </servicios>
  </dia>
  <dia>
    <valor>2008/03/26</valor>
    <servicios>
      <servicio>
        <tipo_servicio>llamada</tipo_servicio>
        <descripcion>
          hora = 08:31
          nro_telefono>3008748523
          duracion>2
        </descripcion>
        <valor>400</valor>
      </servicio>
      <servicio>
        <tipo_servicio>impresion</tipo_servicio>
        <descripcion>
          hora = 09:38
          nro_hojas = 10
        </descripcion>
        <valor>5000</valor>
      </servicio>
    </servicios>
  </dia>
```

```
</detalle_servicios>
```

Seguido al archivo XML creamos el dtd correspondiente:

```
<!ELEMENT detalle_servicios (dia)+>
<!ATTLIST detalle_servicios xmlns CDATA
#FIXED ">
<!ELEMENT dia (valor,servicios)>
<!ATTLIST dia xmlns CDATA #FIXED ">
<!ELEMENT servicios (servicio)+>
<!ATTLIST servicios xmlns CDATA #FIXED ">
<!ELEMENT
servicio
(tipo_servicio,descripcion,valor)>
<!ATTLIST servicio xmlns CDATA #FIXED
">
<!ELEMENT descripcion (#PCDATA)>
<!ATTLIST descripcion xmlns CDATA #FIXED ">
<!ELEMENT valor (#PCDATA)>
<!ATTLIST valor xmlns CDATA #FIXED ">
```

Ahora se crea el XQUERY llamado al XML y realizando las consultas que deseamos:

```
<html>
<body>

<h1>Detalles</h1>
<table border="1">
  <thead>
    <tr>
      <th>
        <td>Dias</td>
        <td>Servicio</td>
        <td>Valor</td>
      </th>
    </thead>
    <tbody>
      <tr>
        <td>
          {for $x at $i in doc("detalle.xml")/detalle_servicios/dia
          let $zero := 0
          let $y := sum(data($x/servicios/servicio/valor))
          return <tr>{
          for $z in $x/servicios/servicio
          return
            <tr><td>{data($x/valor)}</td><td><td>
              colspan="2">{data($z/tipo_servicio)}</td><td>{data($
              z/valor)}</td></tr> }</tr>
          }
          {
            let $total :=
            sum(data(doc("detalle.xml")/detalle_servicios/dia/ser
            vicios/servicio/valor))
```

```

return
<tr><td>Total</td><td>{data($total)}</td></tr>
</table>
</body>
</html>

```

Finalmente construimos el XSL el cual representara el documento XML:

```

<xsl:template match="/detalle_servicios">
  <body>
    <table border="1">
      <tr>
        <td>Dias</td>
        <td>Servicios</td>
        <td>Valor</td>
      </tr>
      <xsl:variable name="total"/>
      <xsl:for-each select="dia">
        <xsl:variable name="p"
select="sum(../servicio/valor)"/>
        <xsl:for-each select="../servicio">
          <tr>
            <td>
<xsl:value-of select="../../valor"/>
            </td>
            <td>
<xsl:value-of select="./tipo_servicio"/>
            </td>
            <td>
<xsl:value-of select="./valor"/>
            </td>
          </tr>
        </xsl:for-each>
      <tr>
        <td>Total Dia</td>
        <td colspan="2">
<xsl:value-of select="sum(../servicio/valor)"/>
        </td>
      </tr>
    </xsl:for-each>
    <tr>
      <td>Total</td>
      <td colspan="2">
<xsl:value-of
select="sum(dia/servicios/servicio/valor)"/>
      </td>
    </tr>
  </table>

```

```

</body>
</xsl:template>

```

Finalmente lo que generamos a partir de todo lo anterior es la siguiente tabla:

Tabla 1. Resultados

Dias	Servicios	Valor
2008/03/25	internet	500
2008/03/25	internet	3000
Total Dia	3500	
2008/03/26	llamada	400
2008/03/26	impresion	5000
Total Dia	5400	
Total	8900	

### V. CONCLUSIONES

Para concluir podemos afirmar que entre las diferentes tecnologías del manejo de datos e información el XML, XSL y XQUERY son las más utilizadas, brindando rapidez y representando un costo reducido y con las ventajas de lograr resultados en corto plazo. El XML lo encontramos desde la creación de los mismos documentos, ayudando a definir una estructura que sea fácilmente entendible por las maquinas. Por esto XML es un metalenguaje que todo profesional en información, ingeniero y archivista debe conocer detalladamente y aplican prácticamente por medio de la utilización de editores y herramientas especializadas que le permiten utilizar toda su potencialidad, de ese aprendizaje dependerá su continuidad y adaptación en el mundo de la información, el cual siempre se encuentra en continua evolución, además prácticamente toda la información y documentos que circulan en las redes tienen algún componente en XML acompañado por XSL y XQUERY.

### VI. REFERENCIAS

- [1] Tim. B. Extensible Markup Language (XML) 1.0. World Wide Web Consortium. 16 de agosto de 2006, [www.w3.org/TR/REG.xml](http://www.w3.org/TR/REG.xml)
- [2] Wikipedia contributions, Wikipedia the free encyclopedia, 27 junio 2011, [http://es.wikipedia.org/wiki/Extensible\\_Stylesheet\\_Language\\_Transformations](http://es.wikipedia.org/wiki/Extensible_Stylesheet_Language_Transformations)
- [3] Brun. R. XML y la gestión de contenidos [online]. "Hipertext.net" num 3, 2005 <<http://www.hipertext.net>> [consulta: 26/06/2007].ISSN 1695-5498
- [4] Bombadil "XSLT (I): hojas de estilos para transformación", 24 agosto 2010, <http://bosqueviejo.net/2010/08/24/xslt-i-hojas-de-estilos-para-transformacion>

- [5] P. Green, “Ejemplo sobre XML y XSL a XHTML”, <http://blog.csd-web.com/ejemplo-sobre-xml-y-xsl-a-xhtml/>, May 17 2010.
- [6] J. Perez Rojas , “XML: XPath – Xquery”, Universidad de Talca, II Semestre 2006. Pag. 26-27.
- [7] Wikipedia contributions, Wikipedia the free encyclopedia *Information system de la Wikipedia en inglés, bajo licencia Creative Commons Atribución Compartir Igual 3.0 y GFDL.*, 13 oct 2011 [http://es.wikipedia.org/wiki/Sistema\\_de\\_informaci%C3%B3n](http://es.wikipedia.org/wiki/Sistema_de_informaci%C3%B3n)
- [8] P. Daniel , “Programación XSL - XSLT - Introducción a XSL – XSLT”, 26 de Diciembre de 2005 <http://www.daniel.prado.name/Programacion-XSL-XSLT.asp?art=133>
- [9] “Aplicaciones basadas en XML”, <http://www.clikear.com/manuales/xml/xml4.aspx>
- [10] S. LECOMTE y T. BOULANGER, “XML practico Bases esenciales, conceptos y casos prácticos”, edición abril 2009, ediciones ENI, Pág. 145,146,147,148