

Asignatura

Lenguajes de marcas y sistemas de gestión de información

UNIDAD 6

Almacenamiento de información



UNIVERSAE
Instituto Superior de FP

Sistemas de almacenamiento de información

Podemos almacenar la información XML de distintas maneras:

Ficheros

Es uno de los métodos más usados aunque no se recomienda puesto que no mantiene los estándares de los otros métodos de almacenamiento.

Bases de datos XML habilitadas o enabled

Este método de almacenamiento se conoce como bases de datos no nativas que transforman en esquemas relacionales los archivos XML.

Bases de datos XML nativas

Se basa en XML y usa el documento XML como unidad de almacenamiento no llega a modificar los documentos por lo que pueden ser recuperados en su estado original sin problemas.

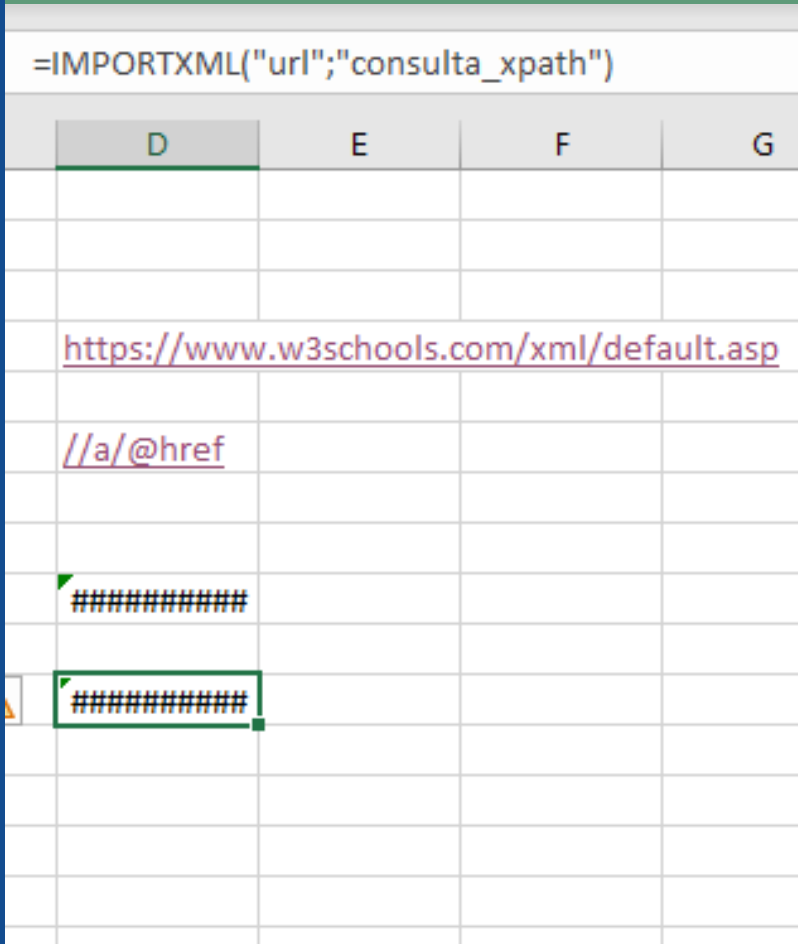




Con esta herramienta podemos importar varios formatos

- La sintaxis es la siguiente:

```
IMPORTXML("url"; "consulta_xpath")
```



Búsqueda de información en documentos XML

Lenguajes de consulta y manipulación



XPath

XPath, un lenguaje path language, con una sintaxis path like que permite buscar y seleccionar nodos en archivos con lenguajes XML



XQuery

XQuery es un lenguaje de consulta y extracción de información y atributos de archivos XML

Los lenguajes de consulta más utilizados son el XPath y el XQuery,

XPath



XPath posee más de 200 comandos predefinidos que podemos manejar para desarrollar la función deseada.

XPath analiza el documento XML con el fin de crear árboles de nodos de siete tipos:

- Nodo raíz • Elemento • Texto • Atributo • Namespace
- Instrucción procesable • Comentario.

Los nodos en Xpath

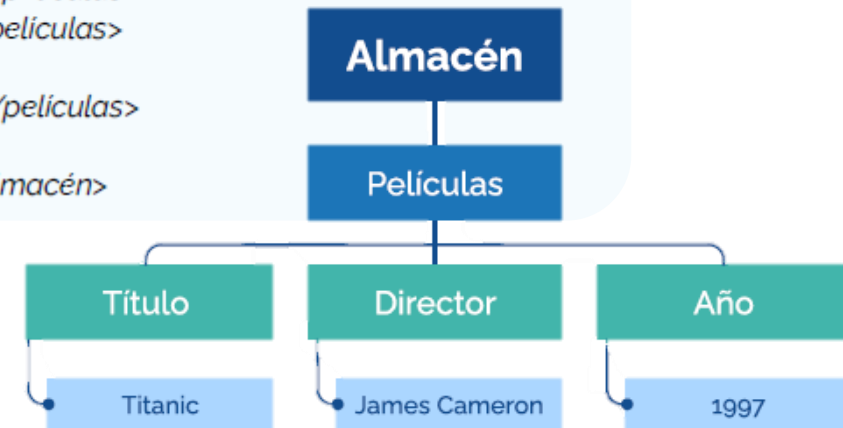
- **Nodo.** Los valores intermedios.
- **Nodo raíz.** El nodo principal es aquel que inicia el archivo.
- **Valor atómico.** Elementos sin padres ni hijos.
- **Ítem o elemento.** Puede ser tanto un nodo como un valor atómico.

Las relaciones entre ellos pueden ser:

- Padre (parent).
- Hijo (children).
- Hermanos (siblings).
- Antecesoros (ancestors)
- Descendientes (descendant).

```
<?xml version="1.0" encoding="UTF-8"?>
```

```
<Almacén>
  <películas>
    <título lang="es">Harry Potter</título>
    <director>J K. Rowling</director>
    <año>2005</año>
  </películas>
  <películas>
    ...
  </películas>
  ...
</Almacén>
```



XPath



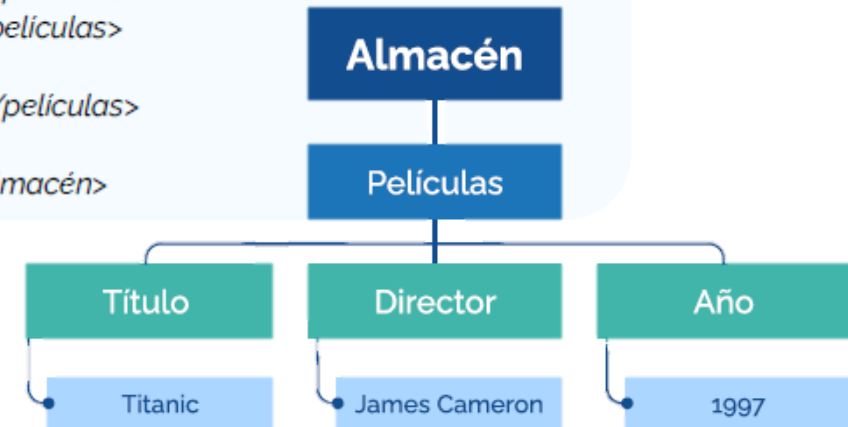
El lenguaje XPath usa expresiones de ruta para navegar por el documento:

Expresiones de ruta	Descripción
Nombre_del_nodo	Selecciona los nodos con el nombre especificado
/	Selecciona el nodo raíz
//	Selecciona los nodos que coinciden con el nodo actual
.	Selecciona el nodo actual
..	Selecciona el nodo padre del actual
@	Selecciona atributos

Expresiones de ruta	Descripción
director	Selecciona los nodos "director"
/Almacén	Selecciona el nodo raíz
//titulo lang="es"	Selecciona los nodos «titulo lang="es"»
//director //año	Selecciona los nodos "director" y "año"
//@lang	Selecciona el atributo "lang"

```
<?xml version="1.0" encoding="UTF-8"?>
```

```
<Almacén>
  <películas>
    <titulo lang="es">Harry Potter</titulo>
    <director>J K. Rowling</director>
    <año>2005</año>
  </películas>
  ...
</películas>
...
</Almacén>
```



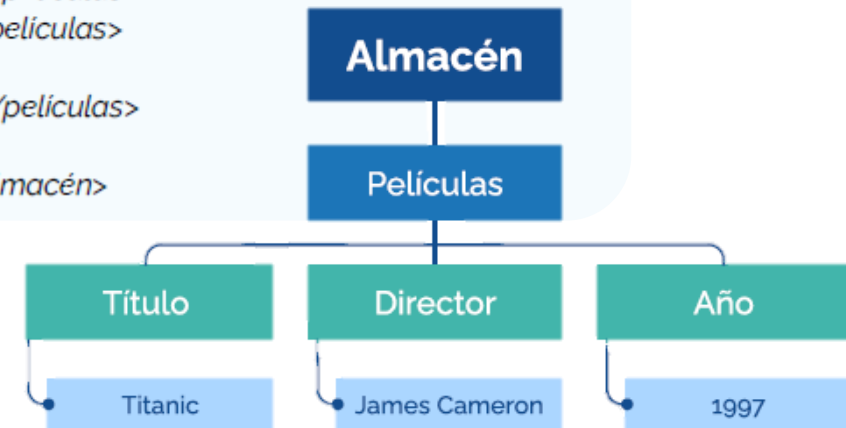
Predicados y operadores

Expresiones	Significado
/Almacén/películas[1]/director	Se selecciona el director del nodo de películas 0, en este caso James Cameron. En ocasiones, dependiendo del navegador, los nodos se empiezan a contar desde el 0.
/Almacén/películas[last()]/director	El director de la última película.
/Almacén/películas[last()-1]/director	El director de la penúltima película.
//título[@lang="es"]	Solo los títulos de películas en español (atributo "es").
/Almacén/películas[año>1990]/director	Los directores de las películas del año 1991 en adelante.
/Almacén/películas[position()<3]/director	Los directores de las dos primeras películas.

Operador	Descripción	Ejemplo
	Computa dos nodos simultáneos	//director //año
+	Adición	1 + 1
-	Sustracción	3 - 1
*	Multiplicación	1 * 3
div	División	4 div 2
=	Igual	año=1997
!=	No igual	año!=1997
<	Menos que	año<1997
<=	Menos o igual que	año<=1997
>	Más que	año>1997
>=	Más o igual que	año>=1997
or	O	año=1997 o año=1996
and	Y	año>1990 y año<2000

```
<?xml version="1.0" encoding="UTF-8"?>
```

```
<Almacén>
  <películas>
    <título lang="es">Harry Potter</título>
    <director>J K. Rowling</director>
    <año>2005</año>
  </películas>
</Almacén>
```



XQuery es un lenguaje de consulta y extracción de información y atributos de archivos XML.
Utiliza las expresiones de XPath, con una sintaxis **FLWOR**.

- **For:** selecciona una secuencia de nodos.
- **Let:** enlaza una secuencia a una variable.
- **Where:** filtra los nodos.
- **Order by:** ordena los nodos.
- **Return:** devolver, lo que nos dará como resultado, se evalúa en todos los nodulos.

```
<html>
<body>
  <hl>
    Almacenaje de filmpedia</hl>
  <ul>
    ! for $x in doc("Inventario21.xml")/Almacén/películas
    where $x/año>1990
    order by $x/título Lang="es"
    return <li>ldata($x/título)</li>
  </ul>
</body>
</html>
```

```
<html>
<body>
  <hl>
    Almacenaje de filmpedia</hl>
  <ul>
    <li>El silencio de los corderos</li>
    <li>El cabo del miedo</li>
    <li>Tacones lejanos</li>
    <li>La bella y la bestia</li>
  </ul>
</body>
</html>
```

Almacenaje de filmpedia:
El silencio de los corderos
El cabo del miedo
Tacones lejanos
La bella y la bestia

Almacenamiento XML nativo

Instalación y ejecución de una base de datos XML nativa:

eXist-db

Requisitos: SO Windows, MAC Os o Linux y Java 1.7 o superior.

Descargaremos la ultima versión de su página oficial.

exist-db.org.

Una vez descargado en la carpeta donde queramos instalarlo lo ejecutaremos con la consola poniendo el siguiente comando.

```
java -jar exist-db-setup-xxx.jar
```

A continuación nos saldrá la interfaz gráfica donde terminaremos la instalación.

ATENCIÓN:

“Nunca debemos instalarlo en “Program Files” o “Archivos de Programa” en ambientes Windows”



Herramientas de tratamiento y almacenamiento de información en formato XML

Existen una multitud de herramientas que nos permiten trabajar con XML y HTML

- Notepad
- Gedit
- editores XML

Para el almacenamiento tenemos eXide, ya que es la usada por eXist-db.

Debemos encontrar las herramientas idóneas a nuestras necesidades.

Estas deben ser similares o no podrían trabajar conjuntamente sin conflictos

The background is a solid blue color. Overlaid on this are several faint, light-blue geometric patterns. A prominent feature is a grid of small squares that forms a stylized world map, with the continents of North and South America visible. Scattered across the entire background are numerous small, light-blue arrows of varying sizes and orientations, some pointing towards the center and others away from it, creating a sense of movement and direction.

UNIVERSAE

— CHANGE YOUR WAY —