

## Unidad 9.

# Conversión y adaptación de documentos XML.

*RA5. Realiza conversiones sobre documentos XML utilizando técnicas y herramientas de procesamiento.*

*Criterios de evaluación:*

- a) Se ha identificado la necesidad de la conversión de documentos XML.*
- b) Se han establecido ámbitos de aplicación.*
- c) Se han analizado las tecnologías implicadas y su modo de funcionamiento.*
- d) Se ha descrito la sintaxis específica utilizada en la conversión y adaptación de documentos XML.*
- e) Se han creado especificaciones de conversión.*
- f) Se han identificado y caracterizado herramientas específicas relacionadas con la conversión de documentos XML.*
- g) Se han realizado conversiones con distintos formatos de salida.*
- h) Se han documentado y depurado las especificaciones.*

**Contenido:**

XPATH. XML Path Language

XSL. eXtensible Stylesheet Language

XSLT

XSL-FO

## XPATH. XML Path Language

### Recursos

[Recomendación w3c](#)

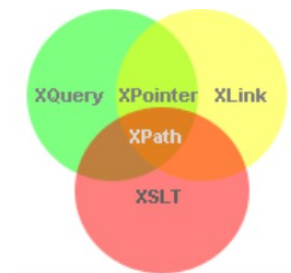
[Material w3schools](#)

XPath es un lenguaje que permite definir expresiones para referenciar los elementos que componen un documento XML. Utiliza **expresiones de ruta** para seleccionar nodos o conjuntos de nodos del documento.

XPath incorpora funciones para flexibilizar el acceso a los elementos de los documentos XML.

Es una recomendación de W3C.

Es un elemento fundamental en otras tecnologías XML, en especial en el estándar XSLT.



XPath considera un documento XML como un árbol de nodos.

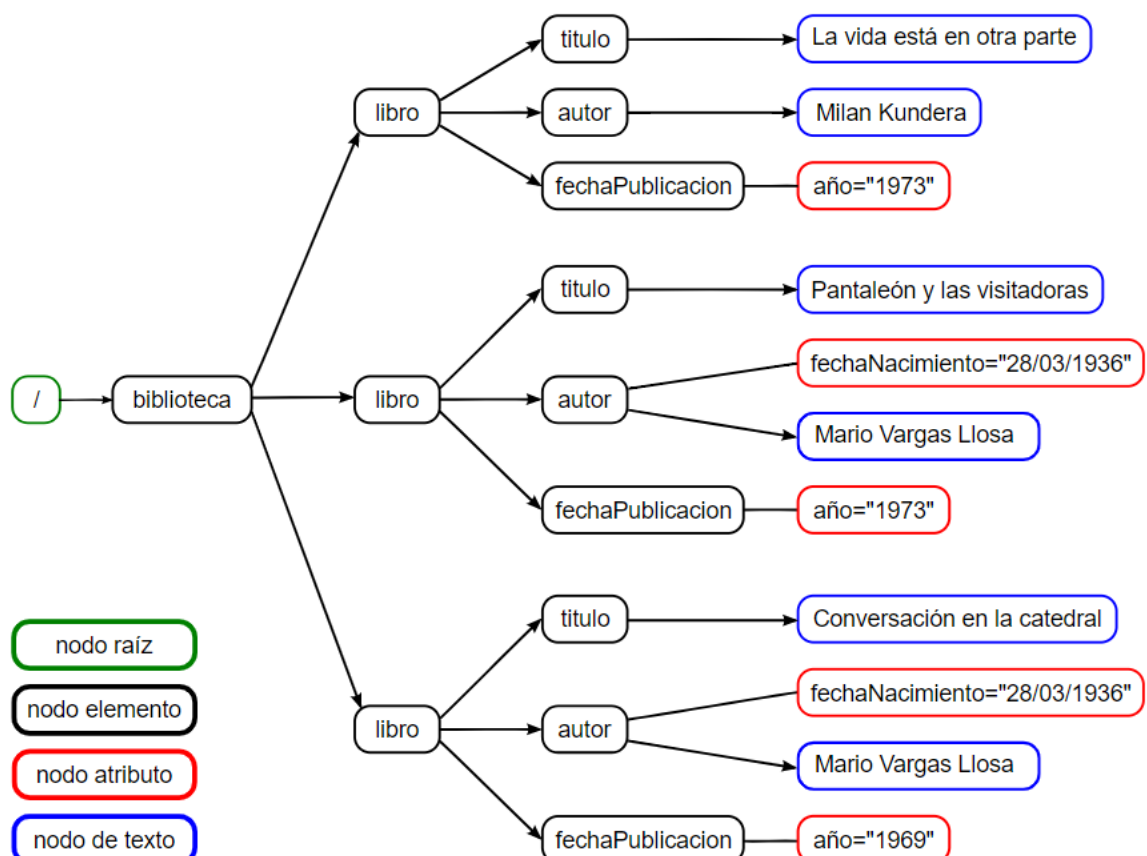
### Nodos

En XPath existen 7 tipos de nodos:

- **Raíz.** Es el elemento / del árbol. No confundir con el elemento raíz del documento XML.
- **Elemento (Element).** Cualquier elemento del documento XML es un nodo elemento del árbol de nodos. Todos ellos tienen un elemento padre, el padre del elemento raíz, es el nodo raíz del documento.  
Pueden tener identificadores únicos, para ello es necesario que un atributo esté definido de ese modo en un DTD o un fichero XSD asociado, esto nos permite referenciar dicho elemento de forma mucho más directa.
- **Atributo (Attribute).** Es un tipo especial de nodo compuesto por nombre y valor. Para las definiciones de los espacios de nombre y para aquellos atributos que se han definido con la propiedad #IMPLIED en su DTD no se crean nodos.
- **Texto (Text).** Es un nodo de los considerados como hoja. Referencia a todos los caracteres que no están delimitados por etiquetas y no tiene hijos.
- Espacio de nombres.
- Instrucciones de procesamiento
- Comentario

## Ejemplo:

```
<?xml version="1.0" encoding="UTF-8"?>
<biblioteca>
  <libro>
    <titulo>La vida está en otra parte</titulo>
    <autor>Milan Kundera</autor>
    <fechaPublicacion año="1973"/>
  </libro>
  <libro>
    <titulo>Pantaleón y las visitadoras</titulo>
    <autor fechaNacimiento="28/03/1936">Mario Vargas Llosa</autor>
    <fechaPublicacion año="1973"/>
  </libro>
  <libro>
    <titulo>Conversación en la catedral</titulo>
    <autor fechaNacimiento="28/03/1936">Mario Vargas Llosa</autor>
    <fechaPublicacion año="1969"/>
  </libro>
</biblioteca>
```



Nodo raíz.	/. No hay que confundirlo con el elemento raíz del documento, en este caso biblioteca
Nodo elemento.	autor
Nodo atributo.	año="1969"
Nodo texto.	Mario Vargas Llosa

### Expresiones XPATH

Una expresión XPath es una cadena de texto que representa un recorrido en el árbol del documento. Las expresiones más simples se parecen a las rutas de los archivos en sistemas operativos.

Para poder evaluar una expresión XPath, el documento debe estar bien formado.

Evaluar una expresión XPath es localizar los nodos del documento que se ajustan al recorrido definido por la expresión.

La evaluación de una expresión Xpath siempre devuelve un conjunto de nodos, aunque puede estar vacío.

Las expresiones Xpath se pueden expresar de dos maneras.

- Expresiones absolutas. Empiezan con el nodo raíz (/) del árbol. El resultado siempre será el mismo, independientemente del contexto actual.
- Expresiones relativas. Se procesan a partir del nodo de contexto.

La evaluación de la expresión XPath se realiza desde el nodo de contexto, que es aquel en el que está situado el proceso. Al inicio del procesamiento el nodo contexto coincide con el raíz e irá cambiando a medida que se van evaluando las expresiones.

Las expresiones XPath se pueden escribir de dos formas distintas:

- Sintaxis abreviada: Compacta y fácil de leer es utilizada en las expresiones más frecuentes.
- Sintaxis completa: Se utiliza una escritura más compleja y permite mayor flexibilidad y un conjunto de opciones más completo.

Sintaxis completa	Sintaxis abreviada
child::	Se omite. Eje por defecto
attribute::	@
descendant-or-self::node()/	//
self::node()	.
parent::node()	..

Las expresiones XPath se pueden dividir en pasos de búsqueda. Cada paso de búsqueda se puede a su vez dividir en tres partes: eje, predicado, selección de nodos.

## Ejes

La navegación dentro de un documento XML a través de XPath tiene lugar desde un nodo de contexto (self) que es el nodo en el que se encuentra el procesador XSLT. El eje nos permite seleccionar un subconjunto de nodos del documento y corresponde a recorridos en el árbol del documento.

Los **nodos elemento** se indican mediante el **nombre del elemento**.

Los **nodos atributo** se indican mediante @ y el nombre del atributo.

El nombre de los ejes define las relaciones de parentesco respecto al nodo de contexto:

- Self. Nodo de contexto.
- Child. Contiene los nodos hijos del nodo de contexto.
- Descendant. Contiene los descendientes del nodo de contexto
- Descendant-or-self. Contiene el nodo de contexto y sus descendientes.
- Parent. Contiene el padre del nodo de contexto.
- Ancestor. Contiene los ancestros del nodo de contexto.
- Ancestor-or-self. Contiene el nodo de contexto y sus ancestros.
- Preceding. Contiene todos los nodos que aparecen antes del nodo de contexto, excluyendo los ancestros.
- Following. Contiene todos los nodos que aparecen después del nodo de contexto, excluyendo los descendientes.
- Preceding-sibling. Contiene los hermanos precedentes del nodo de contexto.
- Following-sibling. Contiene los hermanos situados detrás del nodo de contexto.

**nombre\_nodo.** Selecciona nodos con el nombre nombre\_nodo.

**@nombreatributo.** @ Se utiliza para referenciar nodos atributo.

/ Si está al principio de la expresión, indica el nodo raíz, si no, indica "hijo". Debe ir seguida del nombre de un elemento.

// Descendientes.

.. Nodo padre.

. Nodo hijo.

| Varios recorridos.

## Ejemplos

```
/biblioteca/libro/autor  
/biblioteca//autor  
//@año  
/biblioteca/libro/autor/@fechaNacimiento/..
```

### Predicados.

Establecen una condición a la selección de eje, seleccionando sólo los nodos que cumplen la condición del predicado.

Se escriben entre corchetes.

<b>[@atributo]</b>	Selecciona los elementos <b>que tienen</b> ese atributo.
<b>[n]</b>	Selecciona el elemento de la posición n.
<b>[condicion]</b>	Selecciona los nodos que cumplen la condición. En la expresión condicional se pueden incluir: Operadores lógicos: and, or, not Operadores aritméticos: +,-,*,div,mod Operadores relacionales: =, !=,<,>,<=,>= Las constantes de cadena van entre comillas. La condición puede contener el valor del atributo o el texto que contiene el elemento. Es posible incluir varios predicados seguidos, se aplica en este caso la operación lógica and.

### Selección de nodos.

De los nodos seleccionados por el eje y predicado selecciona los elementos, el texto que contienen o ambos.

*	Cualquier elemento
@*	Cualquier atributo
node()	Cualquier nodo
/node()	Selecciona todos los nodo hijos.
//node()	Selecciona todos los descendientes del nodo.
/text()	Selecciona el texto contenido en el nodo.
//text()	Selecciona el texto contenido en el nodo y en todos sus descendientes.
/*	Selecciona todos los elementos hijos del nodo.
//*	Selecciona todos los elementos descendientes del nodo
/@*	Selecciona todos los atributos del nodo.
//@*	Selecciona todos lo atributos de los descendientes del nodo.
	Operador que permite seleccionar varios caminos (and)

Las expresiones XPath pueden anidarse para lo que permite definir expresiones más complejas.

```
//libro[autor=//libro[titulo="Pantaleón y las visitadoras"]/autor/text()]/titulo
```

En las expresiones Xpath es posible incluir [funciones](#) para conseguir seleccionar un conjunto de nodos mucho más preciso. Algunas de las funciones más importantes son:

<b><i>boolean()</i></b> .	Devuelve true si un conjunto de nodos no es vacío.
<b><i>not()</i></b> .	Al aplicarla sobre un predicado devuelve true si el predicado es falso , y falso si el predicado es true.
<b><i>true()</i></b> .	Devuelve el valor true.
<b><i>false()</i></b> .	Devuelve el valor false.
<b><i>count()</i></b> .	Devuelve número de nodos que forman un conjunto de nodos.
<b><i>name()</i></b> .	Devuelve un nombre de un nodo.
<b><i>local-name()</i></b> .	Devuelve el nombre del nodo actual o del primer nodo de un conjunto de nodos.
<b><i>namespace-uri()</i></b> .	Devuelve el URI del nodo actual o del primer nodo de un conjunto dado.
<b><i>position()</i></b> .	Devuelve la posición de un nodo en su contexto comenzando en 1. Por ejemplo, para seleccionar los dos primeros elementos de tipo elemento de un fichero XML pondremos: <pre>//elemento[position()&lt;=2]</pre>
<b><i>last()</i></b> .	Devuelve el último elemento del conjunto dado.
<b><i>normalize-space()</i></b> .	Permite normalizar los espacios de una cadena de texto, es decir, si se introduce una cadena donde hay varios espacios consecutivos, esta función lo sustituye por uno solo.
<b><i>string()</i></b> .	Convierte un objeto en una cadena. Los valores numéricos se convierten en la cadena que los representa teniendo en cuenta que los positivos pierden el signo. Los valores booleanos se convierten en la cadena que representa su valor, esto es “true” o “false”.
<b><i>concat()</i></b> .	Devuelve dos cadenas de texto concatenadas.
<b><i>string-length()</i></b> .	Devuelve la cantidad de caracteres que forman una cadena de caracteres.
<b><i>sum()</i></b> .	Devuelve la suma de los valores numéricos de cada nodo en un conjunto de nodos determinado.

## [Ejemplo en w3school.](#)

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

<bookstore>

<book category="cooking">
  <title lang="en">Everyday Italian</title>
  <author>Giada De Laurentiis</author>
  <year>2005</year>
  <price>30.00</price>
</book>

<book category="children">
  <title lang="en">Harry Potter</title>
  <author>J K. Rowling</author>
  <year>2005</year>
  <price>29.99</price>
</book>

<book category="web">
  <title lang="en">XQuery Kick Start</title>
  <author>James McGovern</author>
  <author>Per Bothner</author>
  <author>Kurt Cagle</author>
  <author>James Linn</author>
  <author>Vaidyanathan Nagarajan</author>
  <year>2003</year>
  <price>49.99</price>
</book>

<book category="web">
  <title lang="en">Learning XML</title>
  <author>Erik T. Ray</author>
  <year>2003</year>
  <price>39.95</price>
</book>

</bookstore>
```



Escribe las siguientes expresiones Xpath.

Seleccionar todos los nodos de títulos.

Seleccionar el título del primer libro.

Selecciona todos los precios.

Selecciona los nodos con precio superior a 35

Selecciona los nodos de título con precio superior a 35

## Ejemplo.

```
<?xml version="1.0" encoding="UTF-8"?>
<root>
  <employees>
    <employee id="1">Johnny Dapp</employee>
    <employee id="2">Al Pacino</employee>
    <employee id="3">Robert De Niro</employee>
    <employee id="4">Kevin Spacey</employee>
    <employee id="5">Denzel Washington</employee>
  </employees>
  <companies>
    <company id="6">Tata Consultancy Services</company>
    <company id="7">Wipro</company>
    <company id="8">Infosys</company>
    <company id="9">Microsoft</company>
    <company id="10">IBM</company>
    <company id="11">Apple</company>
    <company id="12">Oracle</company>
  </companies>
</root>
```

Escribe las siguientes expresiones XPath

Selecciona el nodo documento.

Selecciona el elemento root.

Selecciona todos los elementos 'employee' que son hijos directos del elemento 'employees'

Selecciona todos elementos 'company' elements independientemente de su posición en el documento.

Selecciona los atributos 'id' de los elementos 'company' independientemente de su posición en el documento.

Selecciona el valor del primer elemento 'employee'

Selecciona el último elemento 'employee'

Selecciona el primer y segundo elemento 'employee', utilizando su posición.

Selecciona todos los elementos 'employee' que tienen el atributo 'id'

Selecciona el elemento 'employee' cuyo valor de atributo 'id' sea 3.

Selecciona todos los elementos 'employee' con 'id' menor o igual a 3.

## Lenguaje Marcas

Selecciona todos los hijos del nodo 'companies'

Selecciona todos los elementos del documento.

Selecciona todos los elementos 'employee' y 'company'

Selecciona el nombre del primer elemento en el documento.

Selecciona el valor numérico del atributo 'id' del primer 'employee'

Selecciona el valor string del atributo 'id' del primer 'employee'.

Selecciona la longitud del contenido del primer elemento 'employee'

Selecciona el número de elementos 'company'

Selecciona la suma de los atributos 'id' del elemento 'company'

## XSL (eXtensible Stylesheet Language)

### Recursos

[Recomendación W3C](#)

[Material w3schools](#)

Es una familia de recomendaciones para la transformación y presentación de documentos XML.

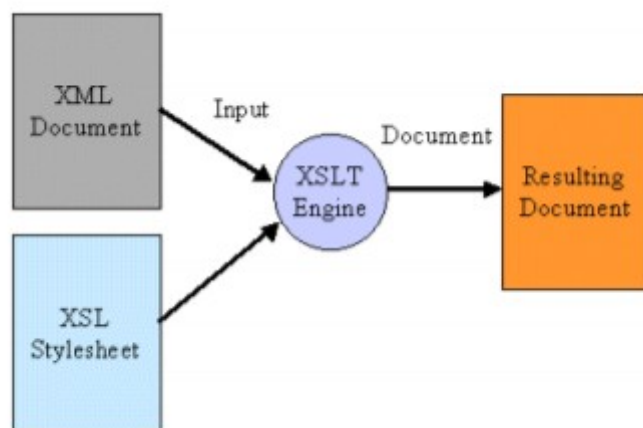
Tecnologías vinculadas:

XSL Transformations. XSLT

XML Path Language. XPATH

XSL Formatting Objects. (XSL-FO)

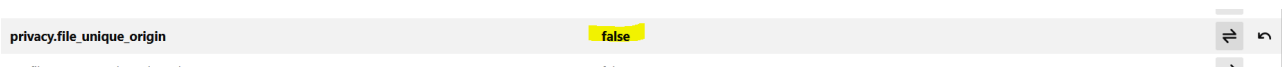
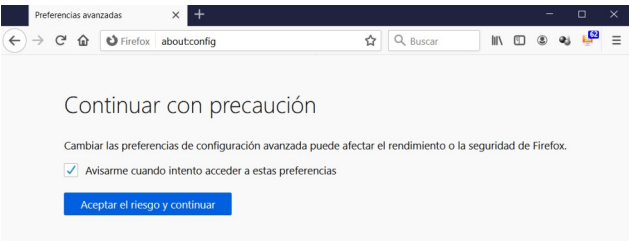
XSL es un lenguaje para expresar hojas de estilo. Una hoja de estilo XSL es, como con CSS , un archivo que describe cómo mostrar un documento XML de un tipo determinado. XSL comparte la funcionalidad y es compatible con CSS2, aunque usa una sintaxis diferente.



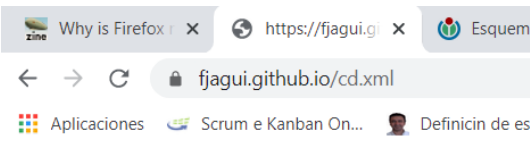
## Lenguaje Marcas

Por motivos de seguridad, el navegador suele tener desactivada la carga de archivos locales. Tenemos dos alternativas para comprobar el funcionamiento de los ejemplos y ejercicios:

Configurar el navegador.



Publicar directamente en github



### My CD Collection

Title	Artist
Empire Burlesque	Bob Dylan
Hide your heart	Bonnie Tyler
Greatest Hits	Dolly Parton
Still got the blues	Gary Moore
Eros	Eros Ramazzotti
One night only	Bee Gees
Sylvias Mother	Dr.Hook
Maggie May	Rod Stewart
Romanza	Andrea Bocelli
When a man loves a woman	Percy Sledge
Black angel	Savage Rose
1999 Grammy Nominees	Many
For the good times	Kenny Rogers
Big Willie style	Will Smith
Tupelo Honey	Van Morrison
Soulsville	Jorn Hoel
The very best of	Cat Stevens
Stop	Sam Brown
Bridge of Spies	T`Pau
Private Dancer	Tina Turner
Midt om natten	Kim Larsen
Pavarotti Gala Concert	Luciano Pavarotti
The dock of the bay	Otis Redding
Picture book	Simply Red
Red	The Communards
Unchain my heart	Joe Cocker

## Ejemplo.

Fichero XML

```
<?xml version="1.0" encoding="UTF-8"?>
<?xml-stylesheet type="text/xsl" href="cd.xsl"?>
<catalog>
  <cd>
    <title>Empire Burlesque</title>
    <artist>Bob Dylan</artist>
    <country>USA</country>
    <company>Columbia</company>
    <price>10.90</price>
    <year>1985</year>
  </cd>
  <cd>
    <title>Hide your heart</title>
    <artist>Bonnie Tyler</artist>
    <country>UK</country>
    <company>CBS Records</company>
    <price>9.90</price>
    <year>1988</year>
  </cd>
  <cd>
    <title>Greatest Hits</title>
    <artist>Dolly Parton</artist>
    <country>USA</country>
    <company>RCA</company>
    <price>9.90</price>
    <year>1982</year>
  </cd>
  <cd>
    <title>Still got the blues</title>
    <artist>Gary Moore</artist>
    <country>UK</country>
    <company>Virgin records</company>
    <price>10.20</price>
    <year>1990</year>
  </cd>
  <cd>
    <title>Eros</title>
    <artist>Eros Ramazzotti</artist>
    <country>EU</country>
    <company>BMG</company>
    <price>9.90</price>
    <year>1997</year>
  </cd>
  <cd>
    <title>One night only</title>
    <artist>Bee Gees</artist>
    <country>UK</country>
    <company>Polydor</company>
    <price>10.90</price>
    <year>1998</year>
  </cd>
  <cd>
    <title>Sylvias Mother</title>
    <artist>Dr.Hook</artist>
    <country>UK</country>
    <company>CBS</company>
    <price>8.10</price>
    <year>1973</year>
  </cd>

```

## Lenguaje Marcas

```
</cd>
<cd>
  <title>Maggie May</title>
  <artist>Rod Stewart</artist>
  <country>UK</country>
  <company>Pickwick</company>
  <price>8.50</price>
  <year>1990</year>
</cd>
<cd>
  <title>Romanza</title>
  <artist>Andrea Bocelli</artist>
  <country>EU</country>
  <company>Polydor</company>
  <price>10.80</price>
  <year>1996</year>
</cd>
<cd>
  <title>When a man loves a woman</title>
  <artist>Percy Sledge</artist>
  <country>USA</country>
  <company>Atlantic</company>
  <price>8.70</price>
  <year>1987</year>
</cd>
<cd>
  <title>Black angel</title>
  <artist>Savage Rose</artist>
  <country>EU</country>
  <company>Mega</company>
  <price>10.90</price>
  <year>1995</year>
</cd>
<cd>
  <title>1999 Grammy Nominees</title>
  <artist>Many</artist>
  <country>USA</country>
  <company>Grammy</company>
  <price>10.20</price>
  <year>1999</year>
</cd>
<cd>
  <title>For the good times</title>
  <artist>Kenny Rogers</artist>
  <country>UK</country>
  <company>Mucik Master</company>
  <price>8.70</price>
  <year>1995</year>
</cd>
<cd>
  <title>Big Willie style</title>
  <artist>Will Smith</artist>
  <country>USA</country>
  <company>Columbia</company>
  <price>9.90</price>
  <year>1997</year>
</cd>
<cd>
  <title>Tupelo Honey</title>
  <artist>Van Morrison</artist>
  <country>UK</country>
  <company>Polydor</company>
```



## Lenguaje Marcas


```
<price>8.20</price>
<year>1971</year>
</cd>
<cd>
  <title>Soulsville</title>
  <artist>Jorn Hoel</artist>
  <country>Norway</country>
  <company>WEA</company>
  <price>7.90</price>
  <year>1996</year>
</cd>
<cd>
  <title>The very best of</title>
  <artist>Cat Stevens</artist>
  <country>UK</country>
  <company>Island</company>
  <price>8.90</price>
  <year>1990</year>
</cd>
<cd>
  <title>Stop</title>
  <artist>Sam Brown</artist>
  <country>UK</country>
  <company>A and M</company>
  <price>8.90</price>
  <year>1988</year>
</cd>
<cd>
  <title>Bridge of Spies</title>
  <artist>T`Pau</artist>
  <country>UK</country>
  <company>Siren</company>
  <price>7.90</price>
  <year>1987</year>
</cd>
<cd>
  <title>Private Dancer</title>
  <artist>Tina Turner</artist>
  <country>UK</country>
  <company>Capitol</company>
  <price>8.90</price>
  <year>1983</year>
</cd>
<cd>
  <title>Midt om natten</title>
  <artist>Kim Larsen</artist>
  <country>EU</country>
  <company>Medley</company>
  <price>7.80</price>
  <year>1983</year>
</cd>
<cd>
  <title>Pavarotti Gala Concert</title>
  <artist>Luciano Pavarotti</artist>
  <country>UK</country>
  <company>DECCA</company>
  <price>9.90</price>
  <year>1991</year>
</cd>
<cd>
  <title>The dock of the bay</title>
  <artist>Otis Redding</artist>
```

## Lenguaje Marcas

```
<country>USA</country>
<company>Stax Records</company>
<price>7.90</price>
<year>1968</year>
</cd>
<cd>
  <title>Picture book</title>
  <artist>Simply Red</artist>
  <country>EU</country>
  <company>Elektra</company>
  <price>7.20</price>
  <year>1985</year>
</cd>
<cd>
  <title>Red</title>
  <artist>The Communards</artist>
  <country>UK</country>
  <company>London</company>
  <price>7.80</price>
  <year>1987</year>
</cd>
<cd>
  <title>Unchain my heart</title>
  <artist>Joe Cocker</artist>
  <country>USA</country>
  <company>EMI</company>
  <price>8.20</price>
  <year>1987</year>
</cd>
</catalog>
```

## Lenguaje Marcas

### Fichero XSL

```
<?xml version="1.0" encoding="UTF-8"?>
<xsl:stylesheet version="1.0" xmlns:xsl="http://www.w3.org/1999/XSL/Transform">
<xsl:template match="/">
<html>
<head>
<style>

    DEFINE ESTILOS
</style>
</head>
<body>
  <h2>My CD Collection</h2>
  <table>
    <tr>
      <th>Title</th>
      <th>Artist</th>
    </tr>
    <xsl:for-each select="catalog/cd">
      <tr>
        <td><xsl:value-of select="title"/></td>
        <td><xsl:value-of select="artist"/></td>
      </tr>
    </xsl:for-each>
  </table>
</body>
</html>
</xsl:template>
</xsl:stylesheet>
```

## XSL-T

### Declaración documento XSLT

```
<?xml version="1.0" encoding="UTF-8"?>
<xsl:stylesheet xmlns:xsl="http://www.w3.org/1999/XSL/Transform" version="1.0">
  <xsl:template match="/">

    </xsl:template>
</xsl:stylesheet>
```

### Declaración equivalente

```
<?xml version="1.0" encoding="UTF-8"?>
<xsl:transform xmlns:xsl="http://www.w3.org/1999/XSL/Transform" version="1.0">
  <xsl:template match="/">

    </xsl:template>
</xsl:stylesheet>
```

### Referencia en el documento XML

```
<?xml version="1.0" encoding="UTF-8"?>
<?xml-stylesheet type="text/xsl" href="..."?>
/.../
```

### Elementos.

#### ***xsl:template***

***<xsl:template match="expresion\_Xpath">***

Se utiliza para crear una plantilla asociada a un elemento XML. Cuando definimos la plantilla como / (xpath) estamos definiendo la plantilla para todo el documento.

```
<xsl:template match="/">
```

#### ***xsl:value-of***

***<xsl:value-of select="expresion\_Xpath"/>***

Recupera el valor del nodo seleccionado. El valor del atributo select es una expresión XPath

```
<xsl:value-of select="title"/>
```

### ***xsl:for-each***

***<xsl:for-each select="expresion\_Xpath">***

Permite realizar un bucle para procesar un conjunto de nodos. Tiene sentido cuando la expresión XPath del atributo select devuelve varios elementos.

### ***xsl:sort select***

***<xsl:sort select = "expresion\_Xpath">***

Elemento utilizado para aplicar un criterio de ordenación en los resultados.

### ***xsl:if***

***<xsl:if test = "expresion\_condicional">***

Elemento que incluye los elementos que se procesaran cuando la expresion condicional sea verdadera.

```
<xsl:if test = "...">  
...  
</xsl:if>
```

### ***xsl:choose***

***<xsl:choose>***

Elemento que permite una selección con múltiples condiciones incluyendo:

***xsl:when test= "expresion\_condicional"***  
***xsl:otherwise***

```
<xsl:choose >  
<xsl:when test= "exp_condicional1">  
<xsl:when test= "exp_condicional2">  
<xsl:otherwise>
```

Actividad propuesta.

- Modifica la plantilla del ejemplo para que sólo se muestren los discos de Bob Dylan
- Modifica la plantilla original para que el lista se ordene por artista.
- Modifica la plantilla original incluyendo precios y posteriormente que sólo aparezcan los discos con precio superior a 10.

## XSL-FO

Un documento XSL-FO es un documento XML en el que se especifica cómo se van a formatear unos datos para presentarlos en pantalla, papel u otros medios. Por ejemplo obtener un documento PDF, documento RTF, PostScript, etc.

Es un estándar de la W3C.

El significado de las siglas XSL-FO es eXtensible Stylesheet Language Formatting Objects. Hay que destacar que en el documento XSL-FO figuran tanto los datos como el formato que se les va a aplicar.

La unidad básica de trabajo en un documento XSL-FO es el "Formating Object", unidad básica para presentar (formatear) la información. Estos objetos de formato se refieren a páginas, párrafos, tablas, etc.

### Estructura de un documento XSL-FO

```
<?xml version="1.0" encoding="utf-8"?>
<fo:root xmlns:fo="http://www.w3.org/1999/XSL/Format">
  <fo:layout-master-set>
    <fo:simple-page-master master-name="A4">
      <!-- Page template goes here -->
    </fo:simple-page-master>
  </fo:layout-master-set>
  <fo:page-sequence master-reference="A4">
    <!-- Page content goes here -->
  </fo:page-sequence>
</fo:root>
```

XSL-FO realiza una división en áreas rectangulares para mostrar la salida:

Páginas.

Todo el contenido se organiza en una o más páginas. (se define el tamaño, los márgenes y las regiones).

Regiones.

Las páginas contienen regiones

Bloques.

Las regiones se componen de bloques.

Cada bloque es un pequeño elementos como: párrafos, tablas, listas, etc.

Líneas.

Son áreas de textos que se incluyen dentro de los bloques.

## Actividades.

- 1 Haz un programa en Python para evaluar expresiones XPath.
- 2 Encuentra las expresiones XPath para los siguientes enunciados:
  - 2.1 Nombre de la Universidad:
  - 2.2 País de la Universidad:
  - 2.3 Nombres de las Carreras:
  - 2.4 Años de plan de estudio de las carreras:
  - 2.5 Nombres de todos los alumnos:
  - 2.6 Identificadores de todas las carreras:
  - 2.7 Datos de la carrera cuyo id es c01:
  - 2.8 Centro en que se estudia de la carrera cuyo id es c02:
  - 2.9 Nombre de las carreras que tengan subdirector:
  - 2.10 Nombre de los alumnos que estén haciendo proyecto:
  - 2.11 Códigos de las carreras en las que hay algún alumno matriculado:
  - 2.12 Apellidos y Nombre de los alumnos con beca:
  - 2.13 Nombre de las asignaturas de la titulación c04:
  - 2.14 Nombre de las asignaturas de segundo trimestre:
  - 2.15 Nombre de las asignaturas que no tienen 4 créditos teóricos:
  - 2.16 Código de la carrera que estudia el último alumno:
  - 2.17 Código de las asignaturas que estudian mujeres:
  - 2.18 Nombre de los alumnos que matriculados en la asignatura a02:
  - 2.19 Códigos de las carreras que estudian los alumnos matriculados en alguna asignatura:
  - 2.20 Apellidos de todos los hombres:
  - 2.21 Nombre de la carrera que estudia Víctor Manuel:
  - 2.22 Nombre de las asignaturas que estudia Luisa:
  - 2.23 Primer apellido de los alumnos matriculados en Ingeniería del Software:
  - 2.24 Nombre de las carreras que estudian los alumnos matriculados en la asignatura Tecnología de los Alimentos:
  - 2.25 Nombre de los alumnos matriculados en carreras que no tienen subdirector:
  - 2.26 Nombre de los alumnos matriculados en asignaturas con 0 créditos prácticos y que estudien la carrera de I.T. Informática :
  - 2.27 Nombre de los alumnos que estudian carreras cuyos planes son anteriores a 2002:



## Fichero XML

```
<universidad>

  <nombre>Universidad de Victoria</nombre>
  <pais>España</pais>

<!-- CARRERAS -->

  <carreras>

    <carrera id="c01">
      <nombre>I.T. Informática</nombre>
      <plan>2003</plan>
      <creditos>250</creditos>
      <centro>Escuela de Informática</centro>
    </carrera>

    <carrera id="c02">
      <nombre>Dipl. Empresariales</nombre>
      <plan>2001</plan>
      <creditos>275</creditos>
      <centro>Facultad de Ciencias Sociales</centro>
    </carrera>

    <carrera id="c03">
      <nombre>Dipl. Relaciones Laborales</nombre>
      <plan>2001</plan>
      <creditos>280</creditos>
      <centro>Facultad de Ciencias Sociales</centro>
      <subdirector>Alfonso Martín Luque</subdirector>
    </carrera>

    <carrera id="c04">
      <nombre>Lic. Química</nombre>
      <plan>2003</plan>
      <creditos>175</creditos>
      <centro>Facultad de Ciencias Experimentales</centro>
    </carrera>

    <carrera id="c05">
      <nombre>Lic. Biología</nombre>
      <plan>2001</plan>
      <creditos>175</creditos>
      <centro>Facultad de Ciencias Experimentales</centro>
    </carrera>

    <carrera id="c06">
      <nombre>Lic. Humanidades</nombre>
      <plan>1980</plan>
      <creditos>475</creditos>
      <centro>Facultad de Humanidades</centro>
    </carrera>

  </carreras>

<!-- ASIGNATURAS -->

  <asignaturas>

    <asignatura id="a01" titulacion="c01">
      <nombre>Ofimática</nombre>
      <creditos_teoricos>3</creditos_teoricos>
      <creditos_practicos>1.5</creditos_practicos>
      <trimestre>1</trimestre>
    </asignatura>

    <asignatura id="a02" titulacion="c01">
      <nombre>Ingeniería del Software</nombre>
      <creditos_teoricos>6</creditos_teoricos>
      <creditos_practicos>1.5</creditos_practicos>
      <trimestre>2</trimestre>
    </asignatura>

    <asignatura id="a03" titulacion="c02">
      <nombre>Administración de Empresas</nombre>
      <creditos_teoricos>4</creditos_teoricos>
```

## Lenguaje Marcas

```
<creditos_practicos>1.5</creditos_practicos>
<trimestre>1</trimestre>
</asignatura>

<asignatura id="a04" titulacion="c02">
  <nombre>Derecho Internacional</nombre>
  <creditos_teoricos>4</creditos_teoricos>
  <creditos_practicos>5</creditos_practicos>
  <trimestre>1</trimestre>
</asignatura>

<asignatura id="a05" titulacion="c04">
  <nombre>Pedagogía</nombre>
  <creditos_teoricos>4</creditos_teoricos>
  <creditos_practicos>1.5</creditos_practicos>
  <trimestre>2</trimestre>
</asignatura>

<asignatura id="a06" titulacion="c03">
  <nombre>Didáctica</nombre>
  <creditos_teoricos>4</creditos_teoricos>
  <creditos_practicos>3</creditos_practicos>
  <trimestre>2</trimestre>
</asignatura>

<asignatura id="a07" titulacion="c04">
  <nombre>Tecnología de los Alimentos</nombre>
  <creditos_teoricos>1.5</creditos_teoricos>
  <creditos_practicos>7.5</creditos_practicos>
  <trimestre>2</trimestre>
</asignatura>

<asignatura id="a08" titulacion="c01">
  <nombre>Bases de Datos</nombre>
  <creditos_teoricos>4.5</creditos_teoricos>
  <creditos_practicos>5.5</creditos_practicos>
  <trimestre>1</trimestre>
</asignatura>

<asignatura id="a09" titulacion="c06">
  <nombre>Historia del Pensamiento</nombre>
  <creditos_teoricos>6</creditos_teoricos>
  <creditos_practicos>0</creditos_practicos>
  <trimestre>2</trimestre>
</asignatura>

</asignaturas>

<!-- ALUMNOS -->

<alumnos>

  <alumno id="e01">
    <apellido1>Rivas</apellido1>
    <apellido2>Santos</apellido2>
    <nombre>Victor Manuel</nombre>
    <sexo>Hombre</sexo>
    <estudios>
      <carrera codigo="c01"/>
      <asignaturas>
        <asignatura codigo="a01"/>
        <asignatura codigo="a03"/>
        <asignatura codigo="a05"/>
        <asignatura codigo="a09"/>
      </asignaturas>
    </estudios>
  </alumno>

  <alumno id="e02">
    <apellido1>Pérez</apellido1>
    <apellido2>García</apellido2>
    <nombre>Luisa</nombre>
    <sexo>Mujer</sexo>
    <estudios>
      <carrera codigo="c02"/>
      <asignaturas>
        <asignatura codigo="a02"/>
        <asignatura codigo="a01"/>
      </asignaturas>
    </estudios>
  </alumno>

</alumnos>
```

## Lenguaje Marcas

```
</asignaturas>
<proyecto>Web de IBM.com</proyecto>
</estudios>

</alumno>

<alumno id="e03" beca="si">
  <apellido1>Pérez</apellido1>
  <apellido2>Romero</apellido2>
  <nombre>Fernando</nombre>
  <sexo>Hombre</sexo>
  <estudios>
    <carrera codigo="c02"/>
    <asignaturas>
      <asignatura codigo="a02"/>
      <asignatura codigo="a01"/>
      <asignatura codigo="a04"/>
      <asignatura codigo="a09"/>
    </asignaturas>
  </estudios>
</alumno>

<alumno id="e04">
  <apellido1>Avalón</apellido1>
  <apellido2>Jiménez</apellido2>
  <nombre>María</nombre>
  <sexo>Mujer</sexo>
  <estudios>
    <carrera codigo="c01"/>
    <asignaturas>
      <asignatura codigo="a02"/>
      <asignatura codigo="a01"/>
      <asignatura codigo="a07"/>
    </asignaturas>
    <proyecto>Estudio de Salinidad del Pantano Iris</proyecto>
  </estudios>
</alumno>

</alumnos>

</universidad>
```

## Lenguaje Marcas

- 3 Resuelve los [ejercicios XPath](#) de mclibre.org
- 4 Resuelve los [ejemplos XPath](#) de w3schools.com

## Actividades

- 1 Crea con Microsoft Office Word un documento (docx). Descomprime el documento (7zip) y haz una breve descripción de los ficheros que lo forman. Localiza donde se encuentra el texto introducido.
- 2 Crea un documento XML que describa el horario del curso. Tiene que tener una etiqueta raíz (por ejemplo, horario), y dentro de ella, etiquetas llamadas lunes, martes, miércoles, jueves y viernes para cada día lectivo. Dentro de cada día habría una etiqueta para cada hora llamadas primera, segunda, tercera, cuarta, quinta y sexta. Cada una de estas etiquetas almacenará el módulo que se imparte en esa hora de ese día.  
Transforma el documento XML mediante XSLT en una tabla HTML, donde aparezca el horario de toda la semana.
- 3 Deseamos utilizar tecnología XML para la gestión académica de los alumnos del Departamento de Informática del IES Gran Capitán.
  - 3.a Diseñar y crear un fichero XML para almacenar la información académica de los alumnos.
  - 3.b Crear un esquema para validar el documento XML.
  - 3.c Crear una página web que permita consultar las notas por cursos.  
Implementa diversos criterios de ordenación.
- 4 Investiga y documenta [XSLT-FO](#) para generar un fichero pdf con tu currículum almacenado en un fichero XML..
- 5 Utiliza el [fichero xml](#) proporcionado por AEMET con la previsión de tiempo para Córdoba. Incluir una plantilla para mostrar la siguiente información.
  - Fecha de la predicción
  - Temperatura máxima para el primer día sobre el que se disponen datos
  - Temperatura mínima para el segundo día sobre el que se disponen datos
  - Periodo/s horarios para el primer día clasificados como "Despejado"
  - Periodo/s horarios para el primer día clasificados que no está "Despejado"
  - Periodo/s horarios para el primer día clasificados como "Muy Nuboso"
  - Periodo/s horarios para el primer día clasificados como que no lloverá.
  - Ultimo Día sobre el que se disponen de datos
  - Temperatura máxima para el penúltimo día sobre el que se disponen datos
  - Temperatura mínima para el ultimo día sobre el que se disponen datos