

Introducción

- Aunque podemos visualizar un fichero XML con un simple editor de texto, no es la manera más amigable ni profesional para presentar los datos que están almacenados en su interior.
- El objetivo de la transformación de documentos XML es extraer información y convertirla para ser representada en otro tipo de documentos (por ejemplo HTML, PDF o RTF).

Transformación - estilos

- Para transformar un documento en cualquier otro tipo de documento disponemos de XSL un estándar del W3C.
 - Suele decirse que XSL es a XML lo que CSS es a HTML.
- XSL permite tomar pleno control sobre los datos, pudiendo establecerse criterios como qué datos ver, en qué orden visualizarlos, estableciendo filtros y definiendo formatos de salida para su representación.

XSL

- *eXtensible Stylesheet language* (lenguaje extensible de hojas de estilo).
 - Permite definir un conjunto de reglas que, aplicado sobre un documento XML, permite transformarlo en un resultado formateado de acuerdo a nuestros intereses.
 - Las reglas serán almacenadas dentro de un fichero llamado hoja de estilo XSL.
- Se compone de tres tecnologías:
 - XSLT – Un lenguaje para la transformación de documentos XML en otros tipos de documentos.
 - XPath – Un lenguaje para navegar por documentos XML y hacer referencia a sus elementos.
 - XSL-FO – Un lenguaje para formatear documentos XML.

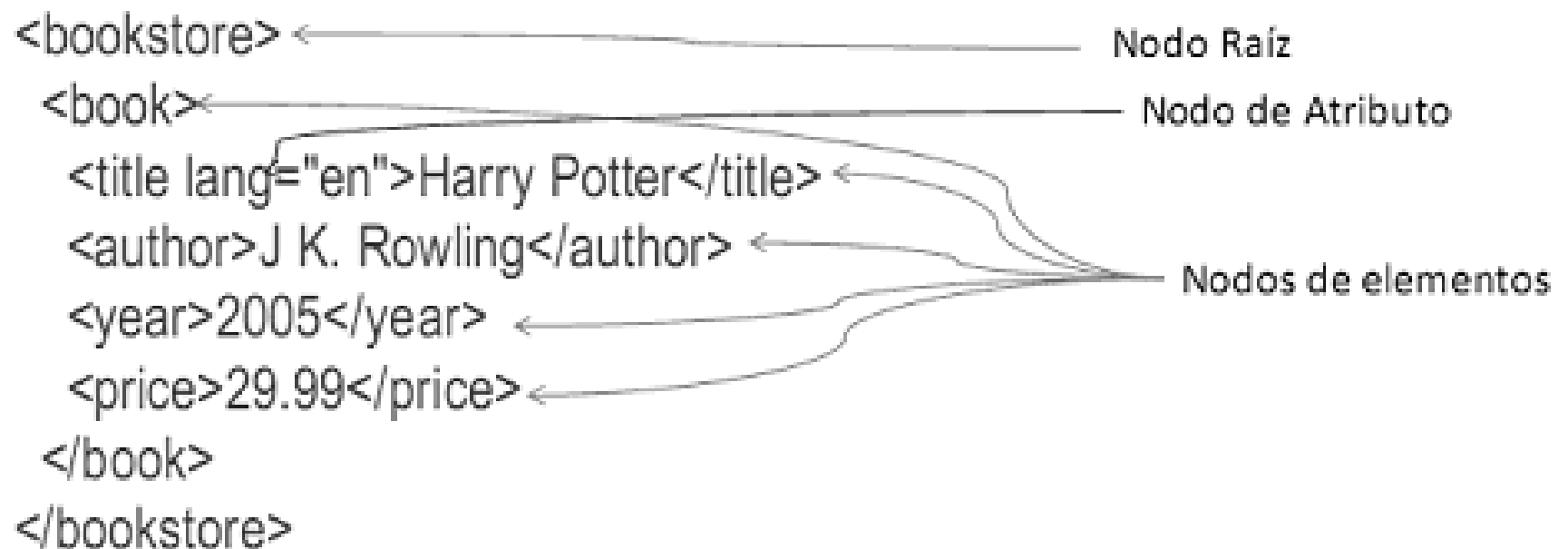
XPATH

- XPath se usa para navegar a través de los elementos y atributos de un documento XML.
 - Es fundamental para otros muchos estándares como XSLT, XQuery, XPointer, XLink y otros.
- XPath es una sintaxis para definir partes de un documento XML. Sus expresiones pueden hacer referencia a un elemento o a un conjunto de elementos de un documento.
- XPath contiene una importante librería de funciones estándar (más de un centenar).

Nodos

- En XPath hay siete tipos de nodos:
 - Elementos.
 - Atributos.
 - Texto.
 - Espacios de nombres.
 - Instrucciones de procesamiento.
 - Comentarios.
 - Nodos de documento.
- Los documentos XML son tratados como un árbol de nodos, donde el elemento contenedor global es llamado “elemento raíz”.
- Los valores atómicos son nodos que no tienen hijos.
- Una expresión XPath es una cadena de texto que representa un recorrido en el árbol del documento.
- Evaluar una expresión XPath es buscar si hay nodos en el documento que se ajustan al recorrido definido en la expresión.

<?xml version="1.0" encoding="ISO-8859-1"?>



Expresiones de ruta

Xpath usa expresiones de ruta para navegar a través de los documentos XML (al estilo de los sistemas de ficheros de los S.O). Las expresiones más utilizadas al describir una ruta son:

Expresión	Descripción
Nombre	Selecciona todos los nodos hijos del nodo nombrado.
/	Selecciona el nodo raíz.
//	Selecciona los nodos del documento desde el nodo actual que cumplen la expresión, independientemente de dónde estén.
.	Selecciona el nodo actual.
..	Selecciona el padre del nodo actual.
@	Selecciona atributos.
	Unión de resultados.

Caracteres comodín

Comodín	Descripción
*	Selecciona cualquier nodo elemento.
@ *	Selecciona cualquier nodo atributo.

Rutas de acceso

- Trata partes de los documentos mediante expresiones de ruta de acceso.
- Cada expresión de ruta es una secuencia de pasos separados por “/”.
- El resultado de una expresión de ruta es un conjunto de nodos.

Acceso a elementos con rutas simples

`/*`

`//@*`

`/fp/modulos/ciclo/titulo`

`//nombre`

`/fp/ciclos/ciclo/@siglas`

`/fp/modulos/modulo/@*`

`//modulo/nombre | //modulo/curso`

Acceso a elementos/atributos vs acceso a su contenido textual

- `/persona/nombre` devuelve el conjunto de nodos llamados “nombre”, descendientes de “persona”, descendiente del nodo raíz. Si no existiera esa ubicación se devolvería el conjunto vacío.
- `/persona/nombre/text()` devuelve el contenido textual del elemento “nombre”.
- `/persona/@dni/data()` devuelve el contenido textual del atributo “dni”.

XPath 2.0: función `string()` para ambos casos.

Operadores

- Aritméticos:

`+, -, *, div, mod`

- Lógicos:

`and, or, not`

- Relacionales:

`=, !=, <, >, <=, >=`

Funciones numéricas

- `round (n)` : Redondeo entero.
- `abs (n)` : Valor absoluto.
- `floor (n)` : Entero inferior.
- `ceiling (n)` : Entero superior.

Funciones de cadena

- `substring(cad, pos1, pos2)`: Subcadena.
- `starts-with(cad, car)`: Verdadero si `cad` empieza por `car`.
- `contains(cad1, cad2)`: Verdadero si `cad2` está contenida en `cad1`.
- `normalize-space(cad)`: Elimina blancos.
- `translate(cad, cad1, cad2)`: Cambia `cad1` por `cad2` en `cad`.
- `string-length(cad)`: Longitud.

Funciones que devuelven elementos por su posición

- `position() = n`: Elemento que se encuentra en la n -ésima posición.
- `elemento[n]`: Ídem.
- `last()`: Último elemento.
- `last() - i`: Último elemento $- i$ (si i es 1, el penúltimo, etc.).

Funciones de agregado

- `count ()` : Conteo de elementos.
- `avg ()` : Media de valores.
- `max ()` : Valor máximo.
- `min ()` : Valor mínimo.
- `sum ()` : Suma de valores.

Predicados

- Los predicados de selección son expresiones usadas para encontrar nodos que contienen valores específicos.
- Pueden seguir cualquier paso en una ruta.
- Se expresan entre corchetes: [Expresión de predicado].

`/banco/cuenta[saldo>400]/titular`

- Se pueden usar operaciones relativas a la posición de los elementos:

`//a[1]`
`//a[position()=2]`
`//a[last()]`

Acceso a elementos con filtros de valores literales

```
/fp/modulos/modulo[@codigo="0373"]
```

```
/fp/ciclos/ciclo[@grado="Medio"]/nombre
```

```
/fp/modulos/modulo/ciclos/ciclo[@siglas="DAW"]/../../duracion
```

```
/fp/ciclos/ciclo[@siglas="ASIR" or @siglas="SMR"]/nombre
```

```
/fp/modulos/modulo[curso="1"][2]
```

```
/fp/ciclos/ciclo[ects][3]
```

```
/fp/modulos/modulo[3 to 5] ← XPath 2.0
```

Acceso a elementos con filtros de valores recuperados

/

```
fp/ciclos/ciclo[@siglas=/fp/modulos/modulo[duracion>=80  
]/ciclos/ciclo/@siglas]
```

/

```
fp/modulos/modulo[number(duracion)>/fp/modulos/modulo[n  
ombre="Sistemas Operativos"]/number(duracion)]/nombre
```