

Introducción a XPath

Willman Acosta Lugo
Profesor

XPath (XML Path Language) es un lenguaje diseñado para navegar y seleccionar nodos en documentos XML. Su importancia radica en la capacidad de acceder de manera eficiente a diferentes partes de un documento XML, lo cual es esencial en el desarrollo de aplicaciones web y multiplataforma. Con este documento trato de proporcionarte una guía detallada sobre XPath, complementada con ejemplos prácticos y referencias académicas para enriquecer el proceso de aprendizaje.

Contenido

1. ¿Qué es XPath y por qué es útil?.....	1
2. Rutas en XPath: Absolutas y Relativas.....	1
3. Filtrado de Nodos con Predicados.....	2
4. Funciones en XPath.....	2
5. Ejemplos Prácticos.....	3
6. Recursos Adicionales.....	5
7. Referencias.....	6



1. ¿Qué es XPath y por qué es útil?

XPath es un lenguaje que permite identificar y navegar por nodos en un documento XML utilizando una sintaxis similar a las rutas de archivos. Es fundamental en tecnologías como XSLT, donde se utiliza para transformar documentos XML. Su utilidad se extiende a la **extracción de información específica, validación de datos y manipulación de estructuras XML** en diversas aplicaciones.[4]

2. Rutas en XPath: Absolutas y Relativas

XPath ofrece dos tipos principales de rutas:

1. **Ruta Absoluta:** Comienza desde la raíz del documento. Por ejemplo, `/biblioteca/libro` selecciona todos los elementos `<libro>` que son hijos directos de `<biblioteca>`. [3]
2. **Ruta Relativa:** Parte del nodo actual en el contexto. Por ejemplo, `libro/titulo` selecciona los elementos `<titulo>` que son hijos de `<libro>` desde el nodo actual.

La elección entre rutas absolutas y relativas **depende del contexto y de la estructura del XML** que se esté manejando. Podríamos, además, conocer otras formas de obtener rutas observando la siguiente tabla:

Path	Descripción
<code>/</code>	Seleccionador del nodo del documento
<code>/raíz</code>	Seleccionador del elemento raíz
<code>/raíz/hijo/nieto/...</code>	Todos los elementos hijos de la raíz
<code>//</code>	Selecciona el nodo, si existe, independientemente de sus posiciones en el documento.
<code>@NombreDelAtributo</code>	Selecciona el atributo correspondiente.
<code>//@id</code>	Selecciona el atributo id, independientemente de su posición en el documento.
<code>//@*</code>	Selecciona cualquier atributo



Path	Descripción
.	Selecciona en el contexto del nodo
..	Selecciona en el padre el contexto del nodo.
*	Selecciona cualquier nodo

Fuente: Autor.

3. Filtrado de Nodos con Predicados

Los predicados en XPath, definidos entre corchetes [], permiten filtrar nodos según condiciones específicas. Por ejemplo:

- `/biblioteca/libro[autor="Gabriel García Márquez"]` selecciona los elementos `<libro>` cuyo autor sea "Gabriel García Márquez".
- `/biblioteca/libro[anio >= 2000]` selecciona los libros publicados desde el año 2000 en adelante.

Los predicados son **herramientas poderosas** para refinar las consultas y obtener resultados más precisos.

4. Funciones en XPath

XPath incorpora funciones que amplían su capacidad de selección y manipulación de datos. Algunas de las más utilizadas incluyen [3]:

- `position()`: Devuelve la posición de un nodo en su contexto.
- `last()`: Selecciona el último nodo en el contexto actual.
- `count()`: Cuenta el número de nodos en un conjunto.
- `contains()`: Determina si una cadena está contenida dentro de otra.
- `starts-with()`: Verifica si una cadena comienza con una subcadena específica.
- `ends-with()`: Comprueba si una cadena termina con una subcadena específica.

Estas funciones permiten construir **consultas más complejas y adaptadas a necesidades específicas**.



5. Ejemplos Prácticos

A continuación, se presentan ejemplos prácticos para ilustrar el uso de XPath en diferentes escenarios. A partir del siguiente XML:

```
<?xml version="1.0" encoding="UTF-8"?>
<biblioteca>
  <libro id="001">
    <titulo>El Quijote</titulo>
    <autor>Miguel de Cervantes</autor>
    <anio>1605</anio>
    <precio>19.95</precio>
  </libro>
  <libro id="002">
    <titulo>Cien Años de Soledad</titulo>
    <autor>Gabriel García Márquez</autor>
    <anio>1967</anio>
    <precio>24.99</precio>
  </libro>
  <libro id="003">
    <titulo>La Sombra del Viento</titulo>
    <autor>Carlos Ruiz Zafón</autor>
    <anio>2001</anio>
    <precio>18.50</precio>
  </libro>
</biblioteca>
```

Tenemos:

a. Seleccionar el último libro de la biblioteca

Utilizando la función `last()`, podemos obtener el último libro:

```
/biblioteca/libro[last()]
```



Unión Europea

SEDE PORVENIR - VIAPOL: C/ Isabela, 1 - 41013 Sevilla SEDE CARTUJA: C/ Leonardo Da Vinci, 7 - 41092 Sevilla

```
▶ /biblioteca/libro[last()]
[14] ✓ Última ejecución: 13:04:38, Duración: 0.2s
... "-/biblioteca[1]/libro[3]"
```

b. Contar libros con precio superior a 20 euros

Para contar los libros cuyo precio es mayor a 20:

```
count(/biblioteca/libro[precio > 20])
```

```
▶ count(/biblioteca/libro[precio > 20])
[15] ✓ Última ejecución: 13:09:56, Duración: 0.2s
... 1
```

c. Obtener títulos de libros cuyo autor termina en "z"

Usando la función `ends-with()`:

```
/biblioteca/libro[ends-with(autor, "z")]/titulo
```

```
//libro[ends-with(autor, "z")]/titulo
[9] ✓ Última ejecución: 12:25:28, Duración: 0.2s
... "-/biblioteca[1]/libro[2]/titulo[1]"
```

d. Seleccionar libros cuyo título contiene la palabra "Sombra"

Utilizando la función `contains()`:

```
▶ //libro[contains(titulo, "Sombra")]/titulo
[10] ✓ Última ejecución: 12:25:28, Duración: 0.2s
... "-/biblioteca[1]/libro[3]/titulo[1]"
```

e. Obtener títulos de libros que no fueron escritos por "Gabriel García Márquez"

Para excluir libros de un autor específico:



`/biblioteca/libro[autor != "Gabriel García Márquez"]/titulo`

```

▶ //libro[autor != "Gabriel García Márquez"]/titulo
[ ]
... [
    "-/biblioteca[1]/libro[1]/titulo[1]",
    "-/biblioteca[1]/libro[3]/titulo[1]"
]

```

f. Otros ejemplos:

Path	Descripción
<code>//libro[1]/texto()</code>	Valor textual del primer elemento 'libro'
<code>//libro[último()]</code>	El último elemento "libro"
<code>//libro[posición() < 3]</code>	Primer y segundo elemento "libro" utilizando su posición
<code>//libro[@id]</code>	Todos los elementos "libro" que tengan un atributo 'id'
<code>//libro[@id='3']</code>	Elemento "título" con el valor del atributo 'id' de '3'
<code>//libro[@id<=3]</code>	Todos los nodos "libro" con el valor del atributo 'id' menor o igual a '3'
<code>/biblioteca/libro/*</code>	Todos los hijos del nodo "libro"
<code>//*</code>	Todos los elementos del documento
<code>//libro //foo:autor</code>	Elementos "libro" y los elementos "autor"
<code>titulo(//*[1])</code>	Nombre del primer elemento del documento
<code>titulo(//libro[1]/@id)</code>	Valor numérico del atributo 'id' del primer elemento "libro"
<code>contar(//libro)</code>	Número de elementos "libro"
<code>suma(//libro/@id)</code>	Suma de los atributos 'id' de los elementos "libro"

Fuente: Autor.

6. Recursos Adicionales

Para profundizar en el estudio de XPath, se recomienda consultar los siguientes recursos:

- **Tutorial de XPath en W3Schools:** Ofrece una introducción detallada y ejemplos prácticos sobre XPath. Disponible en:
https://www.w3schools.com/xml/xpath_intro.asp

- **Tutorial de XPath para principiantes en IONOS:** Proporciona una guía completa sobre la sintaxis y uso de XPath. Disponible en: <https://www.ionos.com/es-us/digitalguide/paginas-web/desarrollo-web/tutorial-de-xpath/>
- **XML Path Language (XPath) 3.1:** Documento de proceso del W3C del 1 de marzo de 2017. Disponible en: <https://www.w3.org/TR/xpath-31/>
- **Introduction to OXPath:** Artículo académico que presenta una extensión de XPath para interactuar con aplicaciones web complejas. Disponible en: https://www.researchgate.net/profile/Ruslan-Fayzrakhmanov/publication/326057256_Introduction_to_OXPath/links/5e3954bf299bf1cdb90b2643/Introduction-to-OPXPath.pdf

7. Referencias

- [1] R. R. Fayzrakhmanov, C. Michels, y M. Neumann, «Introduction to OXPath», arXiv preprint arXiv:1806.10899, 2018.
- [2] «XML Path Language (XPath) 3.1». Disponible en: <https://www.w3.org/TR/xpath-31/>. [Accedido: 11 de marzo de 2025]
- [3] «XPath», Wikipedia, la enciclopedia libre. 26 de abril de 2024. Disponible en: <https://es.wikipedia.org/w/index.php?title=XPath&oldid=159717266>. [Accedido: 11 de marzo de 2025]
- [4] «XPath Tutorial». Disponible en: https://www.w3schools.com/xml/xpath_intro.asp. [Accedido: 11 de marzo de 2025]
- [5] E. R. Harold y W. S. Means, XML in a Nutshell. O'Reilly Media, Inc., 2002.

