

# XPath

Sitio: [Aula Virtual do IES de Teis](#)

Curso: Linguaxes de Marcas e Sistemas de Xestión de Información 2021-22  
(DAM-A)

Libro: XPath

Impreso por: Deivid Durán Durán

Data: Venres, 24 de Xuño de 2022, 00:19

# Táboa de contidos

## 1. XPath

- 1.1. Introducción a XPath
- 1.2. Funciones de XPath
- 1.3. XPath y espacios de nombres

# 1. XPath

(XML PATH Language, lenguaje de búsqueda de datos a través de la jerarquía que les relaciona.)

## XPath

**XPath (XML PATH Language, lenguaje de búsqueda de datos a través de la jerarquía que les relaciona.)** es un estándar aprobado por el **W3C** que **permite navegar entre los elementos y atributos de un documento XML**.

Para hacerlo, se basa en las relaciones de parentesco entre los nodos del documento.

## Expresiones de camino o ruta (*path*)

**XPath** se usa definiendo expresiones de camino, para seleccionar nodos o conjuntos de nodos en un documento **XML**.

Esas expresiones se parecen mucho a las expresiones de camino o ruta (**path**) que se suelen usar en los sistemas de ficheros.

Estas expresiones se aplican a un documento **XML**, asumiendo que su estructura interna es la de un árbol. Al aplicar una expresión, se obtiene un conjunto de nodos (que puede ser vacío).

En el siguiente código **XML** de ejemplo:

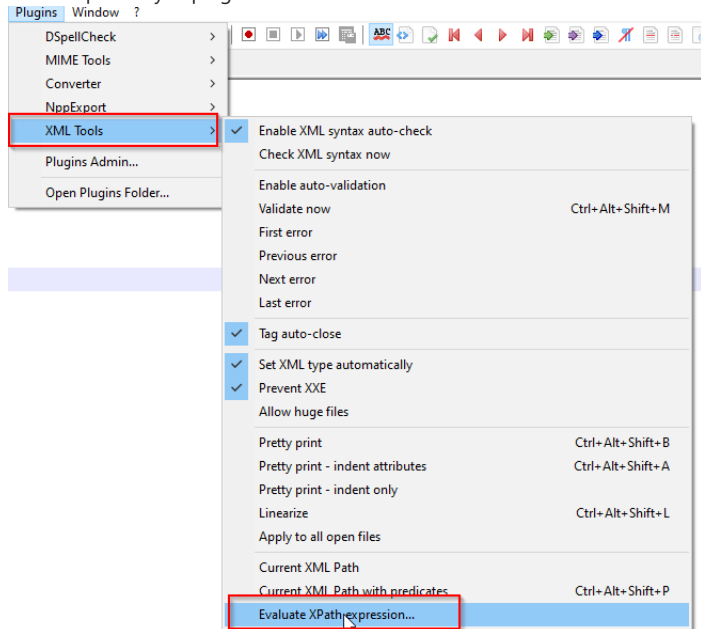
```
<?xml version="1.0" encoding="UTF-8"?>
<matriculas>
  <alumno>
    <nombre>Pedro</nombre>
    <apellido1>López</apellido1>
    <apellido2>Ortega</apellido2>
    <DNI pais="es">López</DNI>
  </alumno>
</matriculas>
```

podemos evaluar la expresión XPath:

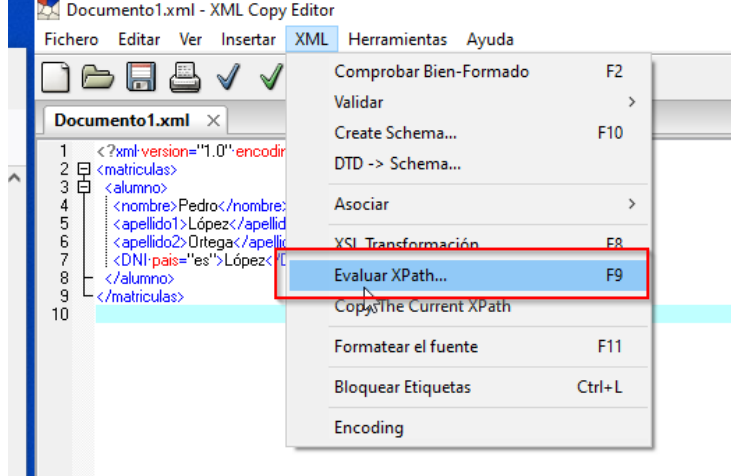
`/matriculas/alumno/nombre`

## Herramientas

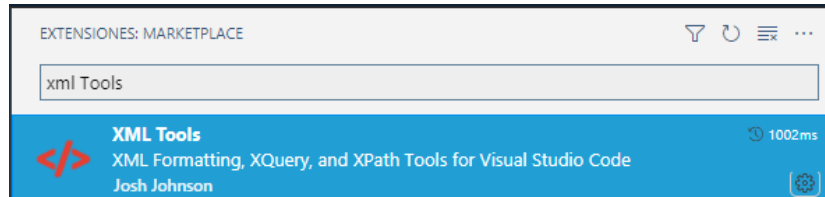
Con Notepad++ y el plugin de XML Tools



Con [XML Copy Editor](#)



Con Visual Studio con la extensión **XML Tools**:



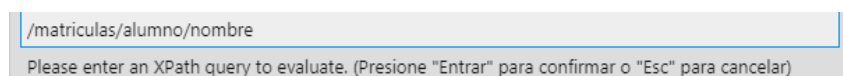
Una vez abierto un documento XML, botón derecho > Paleta de comandos:



Seleccionar la opción **XML Tools: Evaluate XPath** o usar el atajo CTRL+MAYÚS+ALT+X:



A continuación introducir la expresión XPath;



El resultado se abrirá en el terminal. A diferencia de XML Copy Editor, no muestra las etiquetas de apertura y cierre de nombre. En su lugar muestra el nombre del elemento resultado seguido de dos puntos:

En este caso, *nombre*:

XPath Query: /matriculas/alumno/nombre

[Line 4] **nombre:** Pedro



## 1.1. Introducción a XPath

Par ver una introducción a XPath vamos a seguir el enlace disponible en **Recursos** [XPath de McLibre URL](#)

## 1.2. Funciones de XPath

Podemos emplear las funciones para hacer operaciones sobre los conjuntos de elementos o sobre sus contenidos (texto, números).

Entre las funciones más importantes que podemos utilizar en **XPath** destacan:

- **boolean()**, al aplicarla sobre un conjunto de nodos devuelve true si no es vacío. Al aplicarse sobre un número distinto de cero positivo o negativo, devuelve true. Sobre una cadena, devolverá true si no es vacía. False en caso contrario.
- **not()**, al aplicarla sobre un predicado devuelve true si el predicado es falso , y falso si el predicado es true. Por ejemplo: Los autores que no tienen atributo fechaNacimiento: `//autor[not(@fechaNacimiento)]`
- **true()**, devuelve el valor true.
- **false()**, devuelve el valor false.
- **count()**, devuelve el número de nodos que forman un conjunto de nodos. Por ejemplo: Contar el número de elementos libro: `count(//libro)`
- **name()**, devuelve el nombre del primer nodo de un conjunto de nodos.
- **position()**, 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: `//elemento[position()<=2]`
- **last()**, Devuelve el último elemento del conjunto dado. Por ejemplo: `//elemento[last()]`
- **normalize-space()**, 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.
- **string()**, es una función que 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".
- **string-length()**, devuelve la cantidad de caracteres que forman una cadena de caracteres.
- **sum()**, devuelve la suma de los valores numéricos de cada nodo en un conjunto de nodos determinado.
- **starts-with(string s1, string s2)**: devuelve true si s1 empieza por s2, false en caso contrario. Por ejemplo: Los elementos libro cuyo título comienza por La: `//libro[starts-with(titulo, "La")]`
- **contains(string s1, string s2)**: devuelve true si s1 contiene a s2
- **number(str)**, convierte el parámetro a número
- **ceiling(n)** devuelve el menor entero mayor que el parámetro n
- **floor(n)** devuelve el mayor entero menor que el parámetro n
- **round(n)** devuelve el número entero más cercano al parámetro n
- **concat()**, devuelve dos cadenas de texto concatenadas. El ejemplo siguiente devuelve "**XPath** permite obtener datos de un fichero **XML**".

```
concat('XPath', 'permite obtener datos de un fichero XML')
```

Se pueden consultar estas funciones y otras más en el [MDN Web Docs Functions](#)

### 1.3. XPath y espacios de nombres

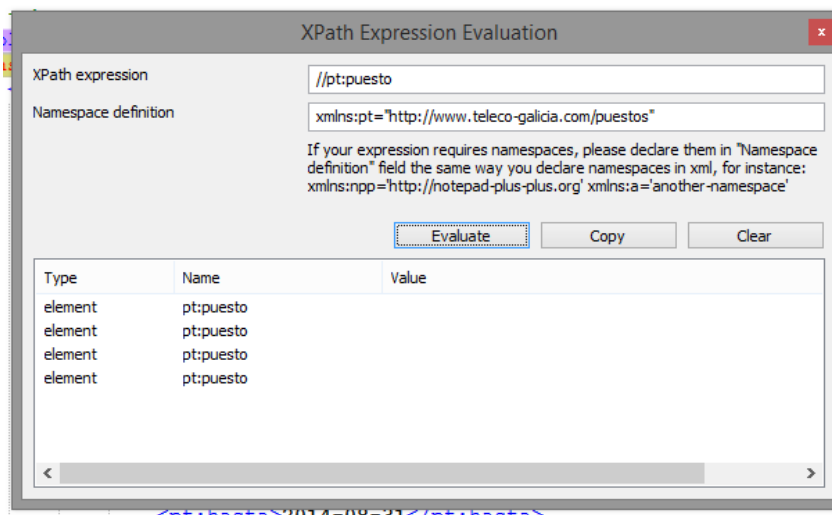
Cuando en un documento XML se definen espacios de nombres, debemos emplear los prefijos que correspondan dentro de las expresiones XPath.

Por ejemplo, en el siguiente documento [empleados-ns-a.xml](#) :

```
1 <?xml version="1.0" encoding="UTF-8" standalone="yes"?>
2 <!-- declaramos un espacio de nombres por defecto que afectará al elemento donde se declara
3      si no tiene un prefijo de otro ns y si no se redefinen otros ns por defecto en su interior
4      Aprovecho para declarar otro ns con prefijo para calificar a los elementos puesto y sus descendientes
5      -->
6 <empleados xmlns="http://www.teleco-galicia.com/empleados"
7      xmlns:pt="http://www.teleco-galicia.com/puestos">
8   <empleado id="0001">
9     <nombre>Juan López Martínez</nombre>
10    <dni>12345678-A</dni>
11    <desde>2015-09-01</desde>
12    <cv>
13      <!-- aquí hay que calificar todos los elementos porque si no se calificaran con el prefijo pt, se les
14           aplicaría el ns por defecto -->
15      <pt:puesto>
16        <pt:nombre>Administrador de sistemas</pt:nombre>
17        <pt:empresa>Telecomunicaciones Galicia S.A.</pt:empresa>
18        <pt:desde>2015-09-01</pt:desde>
19        <pt:hasta></pt:hasta>
20      </pt:puesto>
21      <pt:puesto>
22        <pt:nombre>Soporte de primer nivel</pt:nombre>
23        <pt:empresa>Infraestructuras Norte</pt:empresa>
24        <pt:desde>2015-08-31</pt:desde>
25        <pt:hasta>2014-08-31</pt:hasta>
26      </pt:puesto>
27    </cv>
28  </empleado>
29
30  <empleado id="0002">
31    <nombre>Ana Casal Brea</nombre>
32    <dni>23456789-B</dni>
33    <desde>2021-09-01</desde>
34    <cv>
```

debemos usar el prefijo correspondiente para los elementos correspondientes, por ejemplo: `//pt:puesto`.

Esto es posible en Notepad++ con XML Tools:



Sin embargo, **los espacios de nombre por defecto**, que carecían de prefijo, **no pueden utilizarse en XPath**, por lo que deberemos asociar obligatoriamente un prefijo al espacio de nombre por defecto y prefijar (o calificar) todos los elementos que pertenecían al espacio de nombres por defecto. Por ejemplo: [empleados-ns-b.xml](#) :



```
1 <?xml version="1.0" encoding="UTF-8" standalone="yes"?>
2 <e:empleados xmlns:e="http://www.teleco-galicia.com/empleados"
3 xmlns:pt="http://www.teleco-galicia.com/puestos">
4   <e:empleado id="0001">
5     <e:nombre>Juan López Martínez</e:nombre>
6     <e:dni>12345678-A</e:dni>
7     <e:desde>2015-09-01</e:desde>
8     <e:cv>
9       <pt:puesto>
10        <pt:nombre>Administrador de sistemas</pt:nombre>
11        <pt:empresa>Telecomunicaciones Galicia S.A.</pt:empresa>
12        <pt:desde>2015-09-01</pt:desde>
13        <pt:hasta></pt:hasta>
14      </pt:puesto>
15      <pt:puesto>
16        <pt:nombre>Soporte de primer nivel</pt:nombre>
17        <pt:empresa>Infraestructuras Norte</pt:empresa>
18        <pt:desde>2015-08-31</pt:desde>
19        <pt:hasta>2014-08-31</pt:hasta>
20      </pt:puesto>
21    </e:cv>
22  </e:empleado>
23
24  <e:empleado id="0002">
25    <e:nombre>Ana Casal Brea</e:nombre>
26    <e:dni>23456789-B</e:dni>
27    <e:desde>2021-09-01</e:desde>
28    <e:cv>
29      <pt:puesto>
30        <pt:nombre>Project manager</pt:nombre>
31        <pt:empresa>Telecomunicaciones Galicia S.A.</pt:empresa>
32        <pt:desde>2021-09-01</pt:desde>
33        <pt:hasta></pt:hasta>
34      </pt:puesto>
35      <pt:puesto>
36        <pt:nombre>Consultora funcional</pt:nombre>
37        <pt:empresa>Grupo Norteco</pt:empresa>
38        <pt:desde>2018-02-01</pt:desde>
39        <pt:hasta>2021-08-31</pt:hasta>
40      </pt:puesto>
41    </e:cv>
42  </e:empleado>
43 </e:empleados>
```

Por ejemplo, para buscar las fechas de comienzo del empleado en la compañía usaremos //e:desde y, en Notepad++ con XML Tools habrá que introducir el espacio de nombres.

XPath Expression Evaluation

XPath expression

/e:empleados/e:empleado/e:nombre

Namespace definition

xmlns:e="http://www.teleco-galicia.com/empleados"

If your expression requires namespaces, please declare them in 'Namespace definition' field the same way you declare namespaces in xml, for instance:  
xmlns:npp="http://notepad-plus-plus.org" xmlns:a="another-namespace"

Evaluate

Copy

Clear

Type	Name	Value
element	e:nombre	Juan López Martínez
element	e:nombre	Ana Casal Brea

XML Copy Editor no evalúa expresiones con prefijo de espacio de nombres.

Con Visual Studio y XML Tools sí es posible evaluar expresiones XPath con espacios de nombres:

//e:desde

Please enter an XPath query to evaluate. (Presione "Entrar" para confirmar o "Esc" para cancelar)

Obtendría el siguiente resultado:

PROBLEMAS 1 SALIDA CONSOLA DE DEPURACIÓN TERMINAL

XPath Query: //e:desde

[Line 7] desde: 2015-09-01  
[Line 27] desde: 2021-09-01

Este plugin cuenta con una opción activada por defecto capaz de ignorar los espacios de nombre por defecto. Para que se se tengan en cuenta los espacios de nombres por defecto, deberemos cambiar la configuración:

Archivo > Preferencias > Configuración > Buscamos la cadena xpath y nos aseguramos de que la opción Ignore Default Namespace está deshabilitada:

