



XML: eXtensible Markup Languaje

Tecnologías XML XPath

Dr. Juan Manuel Cueva Lovelle
Departamento de Informática
Universidad de Oviedo
cueva@uniovi.es

Software y estándares para la Web

Esquema

**Grado en
Ingeniería
Informática
del software**

- Introducción a XPath
- XPath
- Conclusiones
- Bibliografía
- Referencias Web
- Ejercicios resueltos

Software y estándares para la Web

Esquema

Grado en
Ingeniería
Informática
del software

- **Introducción a XPath**

- Xpath
- Conclusiones
- Bibliografía
- Referencias Web
- Ejercicios resueltos

Software y estándares para la Web

Introducción a XPath

- **XPath** o XML Path Language: ***sintaxis (no basada en XML) para acceder o referirse a porciones de un documento XML***
- W3C Recommendation 14 December 2010 (*Link errors corrected 3 January 2011; Status updated October 2016*)
 - <https://www.w3.org/TR/xpath20/>

Grado en
Ingeniería
Informática
del software

Software y estándares para la Web

Esquema

Grado en
Ingeniería
Informática
del software

- Introducción a XPath
- **XPath**
- Conclusiones
- Bibliografía
- Referencias Web
- Ejercicios resueltos

Software y estándares para la Web

XPath (I): Objetivo

Grado en
Ingeniería
Informática
del software

- **Objetivo:** Identificar a los elementos del árbol DOM de XML
- Tienen la forma
 - `nodo1/nodo2/.../nodoN`
- Ejemplo: `persona/apellido/materno`
- Describen un camino (*path*)
- Resultado:
 - Un conjunto de nodos
 - String, número o boolean
- Es un estándar del W3C
 - <http://www.w3.org/TR/xpath20/>

Software y estándares para la Web

XPath(II): XPath versus Sistema de Archivos

**Grado en
Ingeniería
Informática
del software**

Archivos y directorios	Nodos dentro de nodos
Respecto a directorio actual	Respecto a nodo actual o de contexto
* = cualquier cosa	* = cualquier nodo
Un archivo por ruta	Uno o varios nodos por ruta

Software y estándares para la Web

XPath (III): Resumen de sintaxis abreviada XPath

Grado en
Ingeniería
Informática
del software

- X – hijo elemento “X”
- * - todos los hijos elemento
- text() - todos los hijos texto
- @Y – atributo “Y”
- X[1] – primer hijo “X”
- X[last()] - último hijo “X”
- */X – nietos “X”
- X//Y – descendientes “Y” de hijo “X”
- //Y – descendientes “Y” de la RAIZ
- //Y/X - descendientes “Y” de HIJO “X”
- .. – padre
- . – todo lo de este nodo
- //X[1][@Y=”Z”] - primeros hijos X con atributo Y=”Z”

Software y estándares para la Web

Esquema

Grado en
Ingeniería
Informática
del software

- Introducción a XPath
- XPath
- **Conclusiones**
- Bibliografía
- Referencias Web
- Ejercicios resueltos

- **XPath** o XML Path Language: *sintaxis (no basada en XML) para acceder o referirse a porciones de un documento XML*
- **W3C Recommendation 14 December 2010 (*Link errors corrected 3 January 2011; Status updated October 2016*)**
 - <https://www.w3.org/TR/xpath20/>

Software y estándares para la Web

Esquema

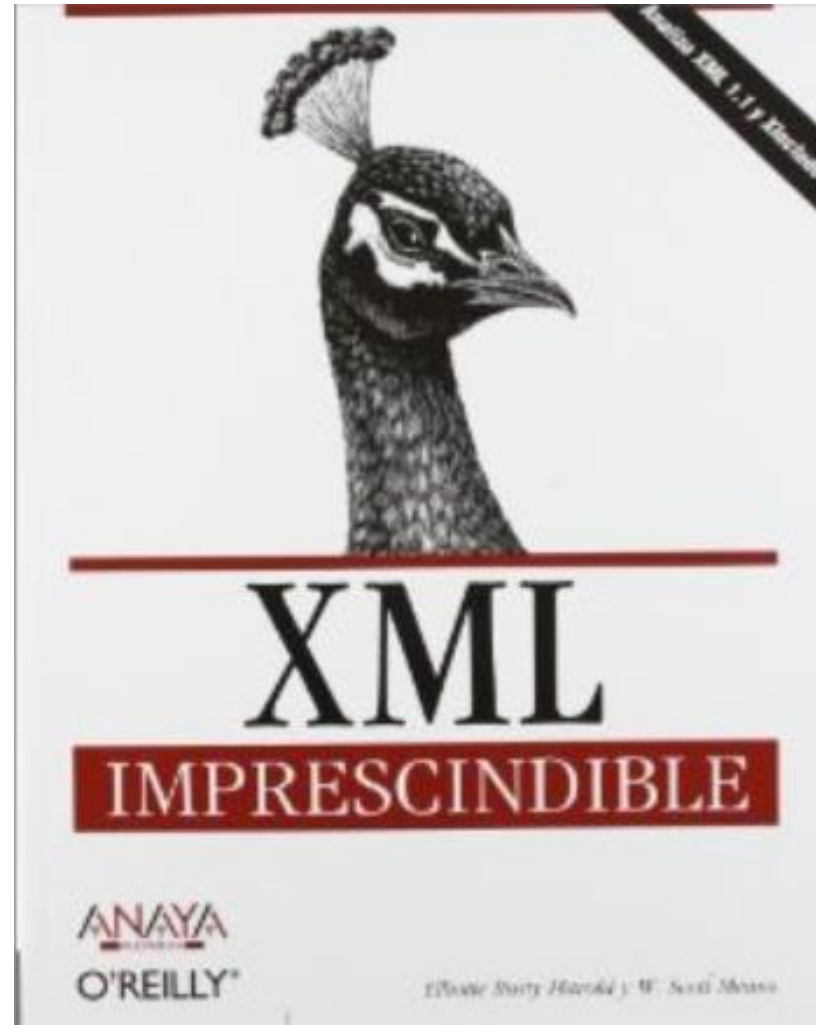
Grado en
Ingeniería
Informática
del software

- Introducción a XPath
- XPath
- Conclusiones
- **Bibliografía**
- Referencias Web
- Ejercicios resueltos

Software y estándares para la Web

Bibliografía (I)

- Libro recomendado de lectura y consulta:
 - “**XML imprescindible**”
 - ANAYA/O'Reilly (2005)
 - E. Rusty Harold y W. Scott Means



Grado en
Ingeniería
Informática
del software

Software y estándares para la Web

Bibliografía (II)

- Libro recomendado de consulta:
 - “**Beginning XML**”
 - John Wiley & Sons (2012)
 - Joe Fawcett, Liam R.E. Quin, and Danny Ayers



Grado en
Ingeniería
Informática
del software

Software y estándares para la Web

Esquema

Grado en
Ingeniería
Informática
del software

- Introducción a XPath
- XPath
- Conclusiones
- Bibliografía
- **Referencias Web**
- Ejercicios resueltos

Software y estándares para la Web

Referencias Web

Grado en
Ingeniería
Informática
del software

- XPath en el W3C
 - <https://www.w3.org/TR/xpath20/>
- Tutorial
 - https://www.w3schools.com/xml/xpath_intro.asp

Software y estándares para la Web

Esquema

Grado en
Ingeniería
Informática
del software

- Introducción a XPath
- XPath
- Conclusiones
- Bibliografía
- Referencias Web
- **Ejercicios resueltos**

- Dado un archivo XML obtener los nodos del árbol DOM de una expresión XPath dada

Software y estándares para la Web

Ejercicios resueltos (II)

Grado en
Ingeniería
Informática
del software

```
# 02010-XPath.py
# # -*- coding: utf-8 -*-
# """
# Ejemplos de uso de XPath
#
# @version 1.2 21/Octubre/2020
# @author: Juan Manuel Cueva Lovelle. Universidad de Oviedo
# """
import xml.etree.ElementTree as ET

def verXPath(archivoXML, expresionXPath):
    """Función verXPath(archivoXML, expresionXPath)
    Visualiza por pantalla el nodo correspondiente de una expresión XPath de un archivo XML

    Version: 1.2 21/Octubre/2020
    Author: Juan Manuel Cueva Lovelle. Universidad de Oviedo
    """
    try:

        arbol = ET.parse(archivoXML)

    except IOError:
        print('No se encuentra el archivo ', archivoXML)
        exit()

    except ET.ParseError:
        print("Error procesando en el archivo XML = ", archivoXML)
        exit()
```

Software y estándares para la Web

Ejercicios resueltos (III)

Grado en
Ingeniería
Informática
del software

```
raiz = arbol.getroot()

# Recorrido de los elementos del árbol
for hijo in raiz.findall(expresionXPath): # Expresión XPath
    print("\nElemento = " , hijo.tag)
    if hijo.text != None:
        print("Contenido = ", hijo.text.strip('\n')) #strip() elimina los '\n' del string
    else:
        print("Contenido = ", hijo.text)
    print("Atributos = ", hijo.attrib)

def main():
    """Prueba de la función verXPath()"""

    print(verXPath.__doc__)

    miArchivoXML = input('Introduzca un archivo XML = ')

    miExpresionXPath = input('Introduzca la expresión XPath = ')

    verXPath(miArchivoXML, miExpresionXPath)

if __name__ == "__main__":
    main()
```

Software y estándares para la Web

Ejercicios resueltos (IV)

Grado en
Ingeniería
Informática
del software

```
C:\WINDOWS\system32\cmd.exe

Función verXPath(archivoXML, expresionXPath)
Visualiza por pantalla el nodo correspondiente de una expresión XPath de un archivo XML

Version: 1.2 21/Octubre/2020
Author: Juan Manuel Cueva Lovelle. Universidad de Oviedo

Introduzca un archivo XML = pizzas.xml
Introduzca la expresión XPath = pizza

Elemento = pizza
Contenido =
Atributos = {'nombre': 'Barbacoa', 'precio': '8.0€'}

Elemento = pizza
Contenido =
Atributos = {'nombre': 'Margarita', 'precio': '6.5€'}

Elemento = pizza
Contenido =
Atributos = {'nombre': 'Tres Quesos con jamón y tocineta', 'precio': '10.5€'}
Presione una tecla para continuar . . .
```

Software y estándares para la Web

Ejercicios resueltos (V)

Grado en Ingeniería Informática del software

```
C:\WINDOWS\system32\cmd.exe
Función verXPath(archivoXML, expresionXPath)
Visualiza por pantalla el nodo correspondiente de una expresión XPath de un archivo XML

Version: 1.2 21/Octubre/2020
Author: Juan Manuel Cueva Lovelle. Universidad de Oviedo

Introduzca un archivo XML = pizzas.xml
Introduzca la expresión XPath = ../

Elemento = pizza
Contenido =
Atributos = {'nombre': 'Barbacoa', 'precio': '8.0€'}

Elemento = ingrediente
Contenido = None
Atributos = {'nombre': 'Salsa Barbacoa'}

Elemento = ingrediente
Contenido = None
Atributos = {'nombre': 'Mozzarella'}

Elemento = ingrediente
Contenido = None
Atributos = {'nombre': 'Tocineta'}

Elemento = ingrediente
Contenido = None
Atributos = {'nombre': 'Ternera'}

Elemento = pizza
Contenido =
Atributos = {'nombre': 'Margarita', 'precio': '6.5€'}

Elemento = ingrediente
Contenido = None
Atributos = {'nombre': 'Tomate'}

Elemento = ingrediente
Contenido = None
Atributos = {'nombre': 'Jamón'}

Elemento = ingrediente
Contenido = None
Atributos = {'nombre': 'Queso'}

Elemento = pizza
Contenido =
Atributos = {'nombre': 'Tres Quesos con jamón y tocineta', 'precio': '10.5€'}

Elemento = ingrediente
Contenido = None
Atributos = {'nombre': 'Cabrales'}

Elemento = ingrediente
Contenido = None
Atributos = {'nombre': 'Mozzarella'}

Elemento = ingrediente
Contenido = None
Atributos = {'nombre': 'Manchego'}

Elemento = ingrediente
Contenido = None
Atributos = {'nombre': 'Jamón'}

Elemento = ingrediente
Contenido = None
Atributos = {'nombre': 'Tocineta'}
Presione una tecla para continuar . . .
```



XML: eXtensible Markup Languaje

Tecnologías XML XPath

Dr. Juan Manuel Cueva Lovelle
Departamento de Informática
Universidad de Oviedo
cueva@uniovi.es