



# Pràctica 8.1: XPath

## Participants:

Santiago Muñoz Cardona  
Ignacio Loayza

## Objectius

Els objectius d'aquesta pràctica d'XPath se centren a desenvolupar una comprensió fonamental de la sintaxi i les funcionalitats bàsiques d'XPath. En primer lloc, caldrà familiaritzar-se amb la sintaxi bàsica de l'XPath i adquirir habilitats per seleccionar elements específics d'un document XML mitjançant camins absoluts i relatius. Un segon objectiu important és la pràctica en la manipulació de text i atributs, que inclourà l'extracció de contingut de text d'elements i l'accés als valors dels atributs.

## Lliuraments

Aquesta pràctica està formada per la part 8.1 i la part 8.2. El lliurament es farà en una sola entrega una vegada ambdues parts estiguin realitzades. Per aquesta raó, de moment no cal lliurar aquesta pràctica.

Els resultats d'aquesta part de la pràctica s'hauran d'entregar en format PDF i l'entrega pot ser a través de GIT\* o el moodle.

**\* S'ha d'entregar l'enllaç del GIT al moodle.**

## Guió

### Exercici 1

Completa la taula explicant els resultats esperats per a cada una de les expressions XPath donades. En casos on la resposta sigui **un objecte sigui substancialment gran**, només cal que indiquis de quin element/s es tracta.

- Si dona **error** o no dona **cap resultat**, explica'n la raó

### Exercici 2

Una vegada tinguis la taula de l'exercici 1 plena, avalua les expressions XPath fent servir el codi Python 3 que es troba a l'arxiu *xpath\_evaluator.py*.

- Assegura't tenir instal·lada la llibreria 'lxml'. Per instal·lar-la, pots executar al terminal: `pip install lxml`

	Ruta XPath	Explica el resultat
1	<code>/llenguatges/llenguatge/nom</code>	<i>Seleccionarà todas las etiquetas nom dentro de llenguatge en la raíz llenguatges. &lt;nom&gt;Kotlin&lt;/nom&gt;</i>
2	<code>/llenguatges/llenguatge/nom/node()</code>	<i>Seleccionarà todos los contenidos de la etiqueta nom dentro de llenguatge en la raíz llenguatges.</i>
3	<code>/llenguatges/llenguatge/paradigmes[@tipat]</code>	<i>Seleccionarà todos los elementos paradigmes que contengan el atributo tipat.</i>
4	<code>/llenguatges/llenguatge/paradigmes[@tipat="false"]</code>	<i>Seleccionarà todos los elementos paradigmes que contengan el atributo tipat con el valor «false».</i>
5	<code>nom</code>	<i>Devuelve la etiqueta nom que no cuelga de un llenguatge</i>
6	<code>/nom</code>	<i>No selecciona nada ya que no se ve la existencia de la raíz.</i>
7	<code>/nom/</code>	<i>No selecciona nada ya que no se ve la existencia de la raíz.</i>
8	<code>//nom</code>	<i>Selecciona todos los elementos nom en el documento sin importar la raíz.</i>
9	<code>//nom/node()</code>	<i>Selecciona todos los contenidos de todas las etiquetas nom en el documento.</i>

10	//llenguatge/nom	<i>Seleccionará todas las etiquetas nom que cuelgan de la etiqueta llenguatge en todo el documento.</i>
11	//llenguatge/nom/node()	<i>Seleccionará todos los contenidos en todas las etiquetas nom que cuelgan de la etiqueta llenguatge en todo el documento.</i>
12	//nom/node()   //popularitat/node()	<i>Seleccionará el contenido en las etiquetas nom y en las etiquetas popularitat en todo el documento.</i>
13	//mode_execucio/*	<i>No puede seleccionar todos los elementos dentro ya que solo contiene node().</i>
14	//llenguatge[2]	<i>Selecciona la segunda etiqueta llamada llenguatge dentro dle programa.</i>
15	//llenguatge[last()-1]	<i>Selecciona la penúltima etiqueta llenguatge.</i>
16	//llenguatge[@fundacio]	<i>Selecciona todas las etiquetas llenguatge que tengan el atributo fundacio.</i>
17	//@fundacio	<i>Selecciona el valor del atributo fundacio en todo el documento.</i>
18	//@fundacio[.>2000]	<i>Selecciona los valores de fundacio mas grandes que 2000 en todo el documento</i>
19	//@fundacio[.>2000]/nom	<i>No puede seleccionar el nombre ya que debe</i>

		<i>volver a la anterior red, fundacio solo contiene un año.</i>
20	<code>//@fundacio[.&gt;2000]/../nom</code>	<i>Selecciona los nombres de llenguatge de todos los que tengan valores de fundacio mas grandes que 2000 en todo el documento</i>
21	<code>//llenguatge[mode_execucio='Java Virtual Machine']/nom/text()</code>	<i>Selecciona el contenido de la etiqueta nom de los llenguatge que tengan el mode_execucio con valor Java Virtual Machine de todo el documento.</i>
22	<code>//llenguatge[nom='Kotlin']/popularitat/node()</code>	<i>Selecciona el contenido de la etiqueta popularitat de llenguatge del cual la etiqueta nom es Kotlin en todo el documento.</i>
23	<code>//mode_execucio[.='Interpretat']/..</code>	<i>Selecciona la etiqueta padre de mode_execucio si es que es igual al valor Interpretat.</i>
24	<code>//paradigmes[node()='Imperatiu']/../nom</code>	<i>Selecciona la etiqueta nom de la etiqueta padre de paradigmes si uno de sus node() contiene Imperatiu</i>
25	<code>//*</code>	<i>Muestra cada elemento y los elementos que lo contienen. Muestra el primer elemento, luego lo que hay dentro luego el siguiente elemento</i>

```
<?xml version="1.0" encoding="UTF-8" ?>
<llenguatges>
  <llenguatge fundacio="2011">
    <nom>Kotlin</nom>
    <mode_execucio>Java Virtual Machine</mode_execucio>
```

```
<popularitat>Creixent</popularitat>
<paradigmes tipat="true">
  <paradigma>Orientat a objectes</paradigma>
  <paradigma>Imperatiu</paradigma>
  <paradigma>Funcional</paradigma>
</paradigmes>
</llenguatge>
<llenguatge fundacio="1983">
  <nom>C++</nom>
  <mode_execucio>Compilació</mode_execucio>
  <popularitat>Alta</popularitat>
  <paradigmes tipat="true">
    <paradigma>Procedimental</paradigma>
    <paradigma>Imperatiu</paradigma>
    <paradigma>Orientat a objectes</paradigma>
    <paradigma>Programació genèrica</paradigma>
  </paradigmes>
</llenguatge>
<llenguatge fundacio="1995">
  <nom>Java</nom>
  <mode_execucio>Java Virtual Machine</mode_execucio>
  <popularitat>Alta</popularitat>
  <paradigmes tipat="true">
    <paradigma>Orientat a objectes</paradigma>
    <paradigma>Basat en classes</paradigma>
  </paradigmes>
</llenguatge>
<llenguatge fundacio="1991">
  <nom>Python</nom>
  <mode_execucio>Interpretat</mode_execucio>
  <popularitat>Molt alta</popularitat>
  <paradigmes tipat="false">
    <paradigma>Imperatiu</paradigma>
    <paradigma>Funcional</paradigma>
  </paradigmes>
</llenguatge>
<llenguatge fundacio="1995">
  <nom>JavaScript</nom>
  <mode_execucio>Interpretat</mode_execucio>
  <popularitat>Alta</popularitat>
  <paradigmes tipat="false">
    <paradigma>Orientat a objectes</paradigma>
    <paradigma>Funcional</paradigma>
  </paradigmes>
</llenguatge>
<nom>Això no és un llenguatge de programació :(</nom>
</llenguatges>
```

## Exercici 3

El fitxer **ods.xml** conté els **Objectius de Desenvolupament Sostenible (ODS)** en català. Els ODS són un conjunt de 17 objectius interconnectats adoptats per les Nacions Unides per abordar els reptes mundials, com la pobresa, la desigualtat, el canvi climàtic i la justícia social, amb l'objectiu de millorar la vida de les persones i protegir el planeta.

<https://www.un.org/sustainabledevelopment/es/objetivos-de-desarrollo-sostenible/>

**Dissenyeu l'expressió XPath** per accedir a la informació sol·licitada al fitxer *ods.xml*. Recordeu que heu de satisfer les condicions de l'enunciat i, tot i conèixer l'XML, heu d'extreure la informació sense afegir cap informació addicional. Per exemple, si se us demana seleccionar l'ODS titulat "Igualtat de Gènere", **heu de seleccionar-lo exclusivament fent servir aquesta informació**. L'exercici seria incorrecte si seleccioneu la informació fent servir el fet que és l'ods=5.

1. El text de dins de l'etiqueta **<titol>** de **tots els ODS** (Objectius de Desenvolupament Sostenible).

Fi de la Pobresa  
Fam Zero  
Salut i Benestar  
Educació de Qualitat  
Igualtat de Gènere  
[...]

```
"//objectiu/titol/node() "
```

2. La descripció (text) de l'ODS titulat **"Igualtat de Gènere"**.

Promou la igualtat de gènere i empoderar totes les dones i nenes. Inclou objectius com eliminar la violència de gènere i garantir la participació igualitària en la presa de decisions.

```
"//objectiu[titol='Igualtat de Gènere']/descripcio/text() "
```

3. Les **accions** (text) de l'ODS que té el títol de **"Fi de la Pobresa"**.

Implementar polítiques socials inclusives  
Garantir protecció social per a tots

```
"//objectiu[titol='Fi de la Pobresa']//accio/node() "
```

4. El **títol** (text) de les ODS **13, 14, 15 i 16**.

Pau, Justícia i Institucions Sòlides

Acció pel Clima  
Vida Submarina  
Vida d'Ecosistemes Terrestres

```
"//objectiu[@ods>=13 and @ods <=16]/titol/node() "
```

5. El **títol** i la **descripció** (text) de l'ODS amb **ods="10"**.

Reducció de les Desigualtats  
Busca reduir les bretxes entre països i dins d'ells. Inclou objectius com empoderar les persones marginades i promoure polítiques inclusives.

```
"//objectiu[@ods=10]/titol/node() |  
//objectiu[@ods=10]/descripcio/node() "
```

6. El llistat d'**accions**, en XML, dels ODS que pertanyen al grup de **tipus econòmic**.

```
<accio>Fomentar l'emprenedoria i la innovació</accio>  
<accio>Reduir la bretxa salarial de gènere</accio>  
[...]
```

```
"grup[@tipus='economic']//accions/accio"
```

7. El text de dins de l'etiqueta **<titol>** del **5è objectiu** dins del grup d'ODS de tipus **"ambiental"**.

Vida Submarina

```
"grup[@tipus='ambiental']/objectiu[5]/titol/text() "
```

8. Dins el grup de tipus **social**, selecciona el **7è objectiu** i retorna la **2a acció**.

```
<accio>Garantir igualtat d'oportunitats per a tothom</accio>
```

```
"grup[@tipus='social']/objectiu[7]/accions/accio[2] "
```

9. El **tipus del grup** al qual pertany l'ODS que té l'acció **"Promoure l'ús d'energies renovables"**.

ambiental

```
'//accio[text()="Promoure l\'ús d\'energiesrenovables"]/../../../../../@tipus'
```

10. Les **5 primeres accions** per les ODS de tipus **ambiental**. [opcional]

```
<accio>Desenvolupar infraestructures per a energies renovables</accio>
```

```
<accio>Promoure l'eficiència energètica a la indústria</accio>  
<accio>Invertir en investigació i desenvolupament</accio>  
<accio>Desenvolupar infraestructures de transport sostenible</accio>  
<accio>Fomentar el reciclatge i la reutilització de productes</accio>
```

No se puede hacer esto en una sola expresión, ya que cuando toma la posición de un elemento, la toma relativamente a su padre, entonces al decir grupo//accion[1] tomará todas las primeras acciones dentro de cada elemento acciones.

Se podría hacer un código en python para contarlas o algo así, pero que pereza.

Pude resolver este problema a la fuerza bruta teniendo en cuenta que por cada objetivo del grupo ambiental, hay 2 acciones.

```
"//grup[@tipus='ambiental']/objectiu[position()<=2]//accio[position()  
<= 2] | //grup[@tipus='ambiental']/objectiu[3]//accio[1]"
```