

# XPath a través de ejemplos

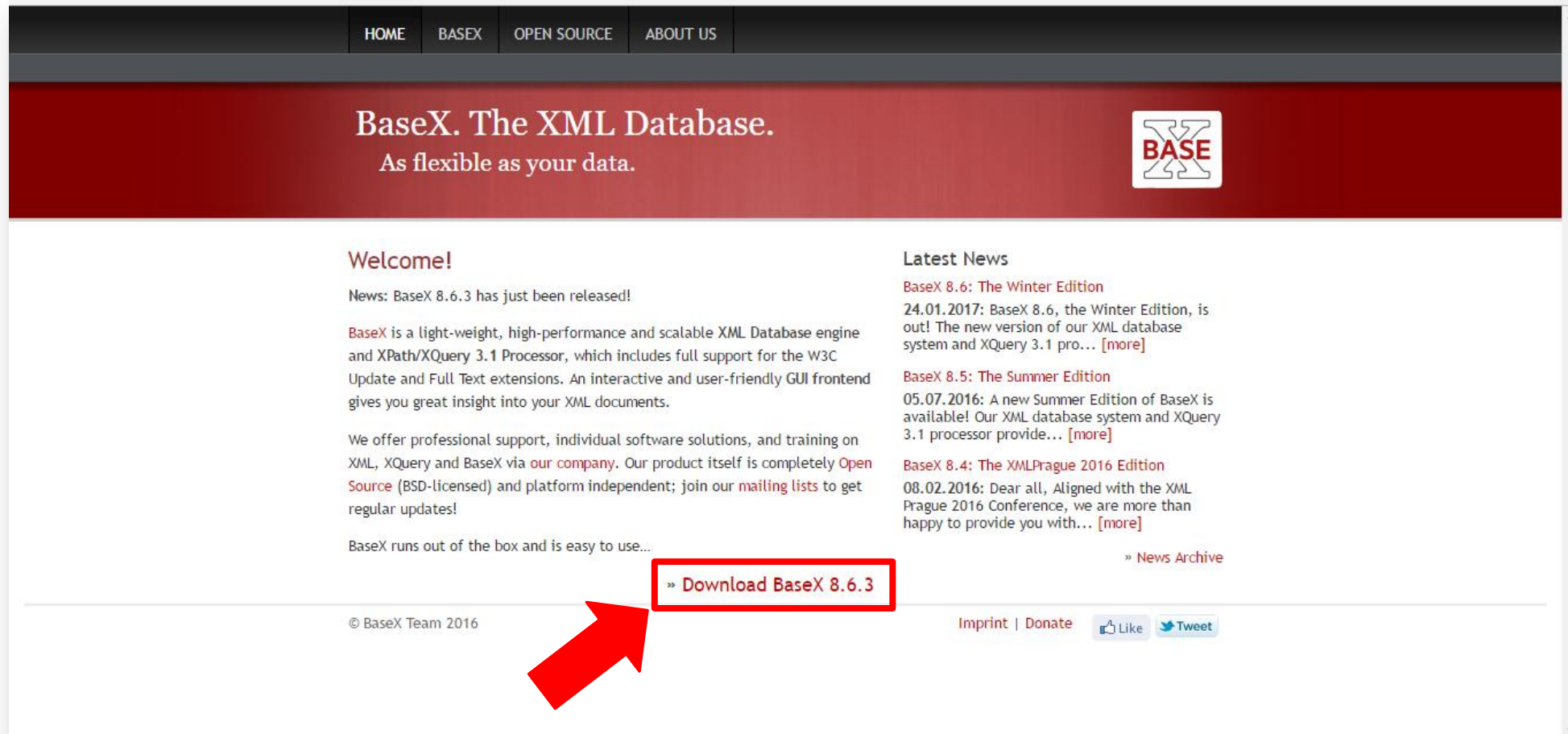
---

PRIMERA PARTE

# Software

## Software necesario para Xpath

- BaseX → Descargar de: <http://basex.org/> e instalar



# Archivo XML

El archivo XML con el que se va a trabajar tiene la siguiente estructura:

1 Crear BD a partir de un XML

2 Seleccionar el archivo XML

3 OK

## Estructura de gim.xml

|           |           |            |          |             |            |           |       |           |            |            |            |           |         |            |         |         |       |       |  |
|-----------|-----------|------------|----------|-------------|------------|-----------|-------|-----------|------------|------------|------------|-----------|---------|------------|---------|---------|-------|-------|--|
| gimnasio  |           |            |          | Educa       |            |           |       |           |            |            |            |           |         |            |         |         |       |       |  |
| gimnasio  |           |            |          |             |            |           |       |           |            |            |            |           |         |            |         |         |       |       |  |
| monitores |           |            |          |             |            |           |       |           |            |            |            |           |         |            |         |         |       |       |  |
| monitor   |           | monitor    |          | actividades |            |           |       | actividad |            |            |            | actividad |         |            |         | monit.  |       |       |  |
| nombre    | telefono  | nombre     |          | nombre      | horario    | sesion    |       | m..       | nombre     | intensid.. | nombre     | horario   | sesion  |            | monit.. |         |       |       |  |
| Mario     | 968222222 | María José |          | Pilates     | sesion     | dia       | dia   | m..       |            | Ligero     | Runing     | sesion    | dia     | horas      | dia     | monit.. |       |       |  |
| apellidos | salario   | apellidos  |          | duracion    | Lunes      | Miércoles |       |           | duracion   | lugar      | duracion   | Lunes     | hora    | Viernes    |         |         |       |       |  |
|           | 1200      |            |          | 55          | horas      | horas     | horas |           | 45         |            | 55         | hora      | 9:00    |            |         |         |       |       |  |
| monitor   |           | telefono   |          | intensidad  | hora       | hora      | hora  | hora      | intensidad |            | intensidad | sesion    | hora    | hora       |         | monit.. |       |       |  |
| nombre    | telefono  | telefono   |          | Ligero      | 8:00       | 16:00     | 9:00  | 16:00     | Ligero     |            | Ligero     | Mié..     | hora    | 9:00       |         |         |       |       |  |
| apellidos | salario   | salario    |          | lugar       |            |           |       |           | horario    |            | lugar      | hora      | 9:00    |            |         |         |       |       |  |
|           | 1300      | 1200       |          | actividad   |            |           |       | Mié..     |            | actividad  |            | nomb..    |         | durac..    |         | horario | mo..  |       |  |
| monitor   |           | monitor    |          | nomb..      | durac..    | horario   |       | monitor.. |            | nomb..     |            | durac..   | horario |            | sesion  |         | m..   |       |  |
| nombre    | telefono  | nombre     | telefo.. | Cro..       | 55         | sesion    | dia   | horas     | monitor    | sesion     | dia        | horas     | dia     |            | dia     |         |       |       |  |
| Anselmo   | 968777..  |            |          |             |            | ..        | hora  | ho..      |            | Jue..      | hora       | 17:00     | horas   |            | horas   |         |       |       |  |
| apellidos | salario   | apellidos  | salario  | intens..    | lugar      |           | 10:00 | 20:00     | monitor    |            |            |           | hora    |            | hora    |         |       |       |  |
|           | 900       |            | 900      | Fue..       | Sala Fit.. |           | hora  | 16:00     |            | monitores  |            | Fue..     |         | Sala Fit.. | 12:00   | 15:00   | 12:00 | 15:00 |  |
|           |           |            |          |             |            |           |       | monitor   |            |            |            |           |         |            |         |         |       |       |  |

# Archivo gym.xml

```
<?xml version="1.0" encoding="UTF-8"?>
<gimnasio>
  <monitores>
    <monitor codigo="m001">
      <nombre>Mario</nombre>
      <apellidos>Cansado Pérez</apellidos>
      <telefono>968222222</telefono>
      <salario>1200</salario>
    </monitor>
    <monitor codigo="m002">
      <nombre>María</nombre>
      <apellidos>Talón Rodríguez</apellidos>
      <telefono>606060606</telefono>
      <salario>1300</salario>
    </monitor>
    <monitor codigo="m003">
      <nombre>María José</nombre>
      <apellidos>Sáez Soriano</apellidos>
      <telefono>699999999</telefono>
      <salario>1200</salario>
    </monitor>
    <monitor codigo="m004">
      <nombre>Anselmo</nombre>
      <apellidos>Solera García</apellidos>
      <telefono>968777777</telefono>
      <salario>900</salario>
    </monitor>
    <monitor codigo="m005">
      <nombre>Barío Son, Calisto</nombre>
      <apellidos>Barío Son</apellidos>
      <telefono></telefono>
      <salario>900</salario>
    </monitor>
  </monitores>

  <actividades>
    <actividad codigo="a10">
      <nombre>Pilates</nombre>
      <duracion>55</duracion>
      <intensidad>Ligero</intensidad>
      <lugar>Sala Fitness</lugar>
      <horario>
        <sesion>
          <día>Lunes</día>
          <horas>
            <hora>8:00</hora>
            <hora>16:00</hora>
          </horas>
        </sesion>
        <sesion>
          <día>Miércoles</día>
          <horas>
            <hora>9:00</hora>
            <hora>16:00</hora>
          </horas>
        </sesion>
      </horario>
      <monitores>
        <monitor codigo="m001"/>
      </monitores>
    </actividad>
    <actividad codigo="a20">
      <nombre>Crosstraining</nombre>
      <duracion>55</duracion>
      <intensidad>Fuerte</intensidad>
      <lugar>Sala Fitness</lugar>
      <horario>
        <sesion>
          <día>Martes</día>
          <horas>
            <hora>10:00</hora>
            <hora>16:00</hora>
            <hora>20:00</hora>
          </horas>
        </sesion>
      </horario>
    </actividad>
    <monitores>
      <monitor codigo="m002"/>
      <monitor codigo="m003"/>
    </monitores>
    <actividad codigo="a30">
      <nombre>Ciclo indoor</nombre>
      <duracion>45</duracion>
      <intensidad>Ligero</intensidad>
      <lugar>Sala ciclo</lugar>
      <horario>
        <sesion>
          <día>Miércoles</día>
          <horas>
            <hora>17:00</hora>
          </horas>
        </sesion>
        <sesion>
          <día>Jueves</día>
          <horas>
            <hora>17:00</hora>
          </horas>
        </sesion>
      </horario>
      <monitores>
        <monitor codigo="m004"/>
      </monitores>
    </actividad>
    <actividad codigo="a40">
      <nombre>Runing</nombre>
      <duracion>55</duracion>
      <intensidad>Ligero</intensidad>
      <lugar>Exterior</lugar>
      <horario>
        <sesion>
          <día>Lunes</día>
          <horas>
            <hora>9:00</hora>
          </horas>
        </sesion>
      </horario>
    </actividad>
    <monitores>
      <monitor codigo="m001"/>
    </monitores>
  </actividades>

  <sesion>
    <día>Miércoles</día>
    <horas>
      <hora>9:00</hora>
    </horas>
  </sesion>
  <sesion>
    <día>Viernes</día>
    <horas>
      <hora>9:00</hora>
    </horas>
  </sesion>
  <monitores>
    <monitor codigo="m004"/>
    <monitor codigo="m005"/>
  </monitores>
  <actividad codigo="a50">
    <nombre>BodyPump</nombre>
    <duracion>55</duracion>
    <intensidad>Fuerte</intensidad>
    <lugar>Sala Fitness</lugar>
    <horario>
      <sesion>
        <día>Martes</día>
        <horas>
          <hora>12:00</hora>
          <hora>15:00</hora>
        </horas>
      </sesion>
      <sesion>
        <día>Jueves</día>
        <horas>
          <hora>12:00</hora>
          <hora>15:00</hora>
        </horas>
      </sesion>
    </horario>
    <monitores>
      <monitor codigo="m001"/>
    </monitores>
  </actividad>
</gimnasio>
```

# XPath en BaseX

2 Ejecutar

file [gim] - BaseX 8.6.3

Database Editor View Visualization Options Help

Find

/gimnasio/monitores

1 Results

1

Escribir la expresión XPath

Editor

1

OK

1 : 1

gimnasio

| monitores |         |         |               | actividades |          |           |         | actividad |          |        |          | actividad |           |        |  |
|-----------|---------|---------|---------------|-------------|----------|-----------|---------|-----------|----------|--------|----------|-----------|-----------|--------|--|
| monitor   | n..     | a..     | te.. sa..     | monitor     | nombre   | horario   | sesion  | m..       | nom..    | inte.. | nombre   | horario   | sesi..    | m..    |  |
| ..        | ..      | ..      | ..            | apellid..   | duraci.. | sesion    | sesion  | m         | dura..   | lugar  | duraci.. | sesion    | día       | sesi.. |  |
| monitor   | n..     | a..     | te.. sa..     | telefono    | intens.. | horas     | horas   | 45        | intens.. | sesion | horas    | hora      | hora      | m..    |  |
| ..        | ..      | ..      | ..            | salario     | lugar    | h.. h..   | h.. h.. | horario   | lugar    | día    | horas    | hora      | hora      | m..    |  |
| monitor   | nom..   | telef.. | nomb.. tele.. | 55          | sesion   | d.. horas | mon..   | sesion    | día      | horas  | mon..    | sesion    | d.. horas | m..    |  |
| apell..   | salar.. | apell.. | sal..         | int..       | lu..     | h.. h..   | h.. h.. | mon..     | int..    | lu..   | h.. h..  | h.. h..   | h.. h..   | m..    |  |
| 900       | 900     |         |               |             |          |           |         |           |          |        |          |           |           |        |  |

monitores

Result

3

Resultado de la expresión XPath

3

XPath

```
<monitores>
  <monitor codigo="m001">
    <nombre>Mario</nombre>
    <apellidos>Cansado Pérez</apellidos>
    <telefono>968222222</telefono>
    <salario>1200</salario>
  </monitor>
  <monitor codigo="m002">
    <nombre>María</nombre>
    <apellidos>Talón Rodríguez</apellidos>
    <telefono>606060606</telefono>
    <salario>1300</salario>
  </monitor>
  <monitor codigo="m003">
    <nombre>María José</nombre>
  </monitor>
</monitores>
```

Query

/gimnasio/monitores

Result:

- Hit(s): 1 Item
- Updated: 0 Items
- Printed: 911 Bytes
- Read Locking: gim
- Write Locking: (none)

Timing:

- Parsing: 0.17 ms
- Compiling: 0.11 ms
- Evaluating: 0.27 ms
- Printing: 0.11 ms
- Total Time: 0.66 ms

# Expresiones simples

## Primera parte

- Obtener los nombres de los monitores
  - `/gimnasio/monitores/monitor/nombre`

```
<nombre>Mario</nombre>  
<nombre>María</nombre>  
<nombre>María José</nombre>  
<nombre>Anselmo</nombre>  
<nombre>Barío Son, Calisto</nombre>
```

- Obtener los nombres de las actividades (sin etiquetas)
  - `/gimnasio/actividades/actividad/nombre/text()`

```
Pilates  
Crosstraining  
Ciclo indoor  
Runing  
BodyPump
```

- Obtener los salario de los monitores.
  - `/gimnasio/monitores/monitor/salario`

```
<salario>1200</salario>  
<salario>1300</salario>  
<salario>1200</salario>  
<salario>900</salario>  
<salario>900</salario>
```

- Obtener el código de las actividades
  - `/gimnasio/actividades/actividad/@codigo`

```
codigo="a10"  
codigo="a20"  
codigo="a30"  
codigo="a40"  
codigo="a50"
```

- Obtener todos los atributos
  - `//@*`

```
codigo="m001"  
codigo="m002"  
codigo="m003"  
codigo="m004"  
codigo="m005"  
codigo="a10"  
codigo="m001"  
codigo="a20"  
codigo="m002"  
codigo="m003"  
codigo="a30"  
codigo="m004"  
codigo="a40"  
codigo="m004"  
codigo="m005"  
codigo="a50"  
codigo="m001"
```

# Operadores

| Operador booleano |               | Operadores aritméticos |                             |
|-------------------|---------------|------------------------|-----------------------------|
| and               | Y lógico      | +                      | Suma                        |
| or                | O lógico      | -                      | Resta                       |
| not               | Negación      | *                      | Producto                    |
| =                 | Igual         | div                    | División                    |
| !=                | Distinto      | mod                    | Resto de la división entera |
| < o &lt;          | Menor         | Otros                  |                             |
| <= o &lt;=        | Menor o igual |                        | Unión                       |
| > o &gt;          | Mayor         |                        |                             |
| >= o &gt;=        | Mayor o igual |                        |                             |

# Haciendo uso de los operadores

- Obtener los nombres de actividades y monitores.

- `/gimnasio/(actividad | monitor)/nombre`

```
<nombre>Mario</nombre>
<nombre>María</nombre>
<nombre>María José</nombre>
<nombre>Anselmo</nombre>
<nombre>Barío Son, Calisto</nombre>
<nombre>Pilates</nombre>
<nombre>Crosstraining</nombre>
<nombre>Ciclo indoor</nombre>
<nombre>Runing</nombre>
<nombre>BodyPump</nombre>
```

- Obtener los monitores que cobran más de 1000 euros

- `/gimnasio/monitores/monitor[salario>1000]/nombre`

```
<nombre>Mario</nombre>
<nombre>María</nombre>
<nombre>María José</nombre>
```

- Obtener el horario de BodyPump

- `/gimnasio/actividades/actividad[nombre='BodyPump']/horario`

- Obtener los códigos de los monitores que dan clase en la sala ciclo y en el exterior.

- `/gimnasio/actividades/actividad[lugar='Sala ciclo' or lugar='Exterior']/monitores/monitor`

```
<horario>
  <sesion>
    <día>Martes</día>
    <horas>
      <hora>12:00</hora>
      <hora>15:00</hora>
    </horas>
  </sesion>
  <sesion>
    <día>Jueves</día>
    <horas>
      <hora>12:00</hora>
      <hora>15:00</hora>
    </horas>
  </sesion>
</horario>
```

```
<monitor codigo="m004"/>
<monitor codigo="m004"/>
<monitor codigo="m005"/>
```



# Funciones de XPath

|           | Función                  | Qué hace   | Ejemplo                                    |
|-----------|--------------------------|--|--|
| Numéricas | <b>round()</b>           | Redondea   | round(7,18) → 7                            |
|           | <b>abs()</b>             | Devuelve el valor absoluto   | abs(-7) → 7                                |
|           | <b>floor()</b>           | Suelo. Entero menor  | floor(7.3) = 7                             |
|           | <b>ceiling()</b>         | Techo. Entero mayor  | ceiling(7.3)                               |
| De cadena | <b>substring()</b>       | Subcadena que comienza en el primer parámetro y tiene, de longitud, el segundo | substring('Beatles', 1,4) → 'Beat'         |
|           | <b>starts-with()</b>     | True, si la cadena comienza por el segundo parámetro                           | starts-with('XML','X') → true              |
|           | <b>contains()</b>        | True, si la cadena contiene el segundo   | contains('XML','XM') → true                |
|           | <b>normalize-space()</b> | Espacios normalizados. Elimina los del principio y final de la cadena          | normalize-space(' The end ') → 'The end'   |
|           | <b>translate()</b>       | Sustituye, en una cadena, la cadena del segundo parámetro, por el tercero      | translate('The end','The','An') → 'An end' |
|           | <b>string-length()</b>   | Longitud de una cadena   | string-length('Beatles') → 7               |

# Expresiones con funciones

---

- Apellidos de los monitores cuyo nombre empiece por 'M'
  - `/gimnasio/monitores/monitor[starts-with(nombre,'M')]/apellidos`

```
<apellidos>Cansado Pérez</apellidos>  
<apellidos>Talón Rodríguez</apellidos>  
<apellidos>Sáez Soriano</apellidos>
```

- Obtener la inicial del nombre de las actividades
  - `/gimnasio/actividades/actividad/substring(nombre,1,1)`

```
P  
C  
C  
R  
B
```

# Funciones por posición

| Función                    | Qué hace   | Ejemplo      |
|----------------------------|--|--------------|
| <b>position()=n</b>        | Elemento que se encuentra en la n-ésima posición   | position()=2 |
| <b>&lt;elemento&gt;[n]</b> | Elemento que se encuentra en la n-ésima posición   | monitor[2]   |
| <b>last()</b>              | El último elemento de un conjunto                  | last()       |
| <b>last() - i</b>          | El último elemento - i (ej. Si i=1 → el penúltimo) | last()-2     |

- Obtener todos los datos de la penúltima actividad
  - `/gimnasio/actividades/actividad[last()-1]`
- Obtener el nombre del tercer monitor
  - `/gimnasio/monitores/monitor[3]/nombre`
  - `/gimnasio/monitores/monitor[position()=3]/nombre`

```
<nombre>María José</nombre>
```

```
<actividad codigo="a40">
  <nombre>Runing</nombre>
  <duracion>55</duracion>
  <intensidad>Ligero</intensidad>
  <lugar>Exterior</lugar>
  <horario>
    <sesion>
      <día>Lunes</día>
      <horas>
        <hora>9:00</hora>
      </horas>
    </sesion>
    <sesion>
      <día>Miércoles</día>
      <horas>
        <hora>9:00</hora>
      </horas>
    </sesion>
    <sesion>
      <día>Viernes</día>
      <horas>
        <hora>9:00</hora>
      </horas>
    </sesion>
  </horario>
  <monitores>
    <monitor codigo="m004"/>
    <monitor codigo="m005"/>
  </monitores>
</actividad>
```