

XPath

XML dokumentu-zuhaitzetan (hel)bideak zehazteko lengoia

XLink, XPointer, XQuery

Xabier Arregi Iparragirre

Xabier Artola Zubillaga

XPath

- XML dokumentuetako guneak eta atalak zehazteko (erreferentziatzeko) lengoaia bat
- W3C XSLT estandarraren osagai funtsezkoa
- W3Cren *XQuery*-n eta *XPointer*-en ere erabilia

Aurkezpenaren edukia

- Sarrera
- Sintaxia
- Kokapen-espresioak
- Espresio komunak
- Funtzioak
- *XLink, XPointer, XQuery*: zer diren eta zertarako erabiltzen diren
- Ariketak

Zer da XPath?

- XPath: XML dokumentu baten atalak zehaztu eta erreferentziatzeko sintaxi bat
- Zuhaitzeko *path* edo "bideen" bitartez erreferentziatuko dira XML elementuak
- Funtzio estandarren multzo bat
- XSLTren funtsezko osagaia
- XPath W3Cren estandar bat da
- XPath ez da XMLz idazten!

XML zuhaitzeko bideak

- *Path* edo bide-izen horiek direktorioetako fitxategiak erreferentziatzeko erabiltzen direnen antzekoak dira:

C:/Dokumentuak/aurkezpena.doc

catalog/cd/title

Adibidea: /catalog

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<catalog>
  <cd country="USA">
    <title>Empire Burlesque</title>
    <artist>Bob Dylan</artist>
    <price>10.90</price>
  </cd>
  <cd country="UK">
    <title>Hide your heart</title>
    <artist>Bonnie Tyler</artist>
    <price>9.90</price>
  </cd>
  <cd country="USA">
    <title>Greatest Hits</title>
    <artist>Dolly Parton</artist>
    <price>9.90</price>
  </cd>
</catalog>
```

Adibidea: /catalog/cd

```
<?xml version="1.0" encoding="ISO-8859-1"?>
```

```
<catalog>
```

```
<cd country="USA">  
  <title>Empire Burlesque</title>  
  <artist>Bob Dylan</artist>  
  <price>10.90</price>  
</cd>
```

```
<cd country="UK">  
  <title>Hide your heart</title>  
  <artist>Bonnie Tyler</artist>  
  <price>9.90</price>  
</cd>
```

```
<cd country="USA">  
  <title>Greatest Hits</title>  
  <artist>Dolly Parton</artist>  
  <price>9.90</price>  
</cd>
```

```
</catalog>
```

Adibidea: /catalog/cd/price

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<catalog>
  <cd country="USA">
    <title>Empire Burlesque</title>
    <artist>Bob Dylan</artist>
    <price>10.90</price>
  </cd>
  <cd country="UK">
    <title>Hide your heart</title>
    <artist>Bonnie Tyler</artist>
    <price>9.90</price>
  </cd>
  <cd country="USA">
    <title>Greatest Hits</title>
    <artist>Dolly Parton</artist>
    <price>9.90</price>
  </cd>
</catalog>
```


Adibidea: funtzioak

```
/catalog/cd[price>10.80]
```

- salneurria 10.80 baino handiagoa duten CD guztiak erreferentziatzen ditu

Sintaxia: adabegien kokapena

- XML dokumentua: adabegiz osaturiko zuhaitza
- XPath patroia (*pattern*): adabegi-zerrenda bat, non osagaiak barra okerrez (*slash-ez*) banaturik baitaude:
 - ◆ **/catalog/cd/price**
bide absolutua: catalog-en ume diren cd elementuen price elementu guztiak
 - ◆ **//cd**
cd elementu guztiak, edozein mailatan daudela ere

Adibidea

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<catalog>
  <cd country="USA">
    <title>Empire Burlesque</title>
    <artist>Bob Dylan</artist>
    <price>10.90</price>
  </cd>
  <cd country="UK">
    <title>Hide your heart</title>
    <artist>Bonnie Tyler</artist>
    <price>9.90</price>
  </cd>
  <cd country="USA">
    <title>Greatest Hits</title>
    <artist>Dolly Parton</artist>
    <price>9.90</price>
  </cd>
</catalog>
```

Sintaxia: elementu ezezagunak, izartxoaren bidez (*)

- `/catalog/cd/*`
catalog-en ume diren cd elementu guztien ume guztiak
- `/catalog/*/price`
catalog elementuaren umeen ume diren price elementu guztiak
- `/**/price`
bi arbaso dituzten price elementu guztiak
- `/**`
dokumentuko elementu guztiak

Sintaxia: adarrak zehaztu (espezifikatu), predikatuen (baldintzen) bitartez: []

- `/catalog/cd[1]`
catalog-en lehen cd umea
- `/catalog/cd[last()]`
catalog-en azken cd umea
- `/catalog/cd[price]`
catalog-en ume diren cd guztietarik, price elementu umeren bat dutenak
- `/catalog/cd[price=10.90]`
price-ren baliozat 10.90 dutenak soilik (cd elementuak!)
- `/catalog/cd[price=10.90]/price`
catalog-en ume diren cd-en price elementu guztiak, baldin eta price horren balioa 10.90 bada

Sintaxia: hainbat bide patroi berean: |

- `/catalog/cd/title | /catalog/cd/artist`
catalog-en ume diren cd guztien title eta artist elementu ume guztiak
- `//title | //artist`
dokumentuko title eta artist elementu oro
- `//title | //artist | //price`
dokumentuko title, artist eta price elementu oro
- `/catalog/cd/title | //artist`
catalog-en ume diren cd guztien title ume guztiak eta dokumentuko artist elementu oro

Adibidea

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<catalog>
  <cd country="USA">
    <title>Empire Burlesque</title>
    <artist>Bob Dylan</artist>
    <price>10.90</price>
  </cd>
  <cd country="UK">
    <title>Hide your heart</title>
    <artist>Bonnie Tyler</artist>
    <price>9.90</price>
  </cd>
  <cd country="USA">
    <title>Greatest Hits</title>
    <artist>Dolly Parton</artist>
    <price>9.90</price>
  </cd>
</catalog>
```

Sintaxia: atributuak patroietan: @

- **//@country**
country izeneko atributu oro dokumentu osoan
- **//cd[@country]**
country izeneko atributua duten cd elementu guztiak
- **//cd[@*]**
atributuren bat duten cd elementu guztiak
- **//cd[@country='UK']**
UK balioa duen country izeneko atributua duten cd elementu guztiak

Adabegiak adierazteko bideak: kokapen-bideak (*location path*)

- Bide-espresio batek adabegi-multzo (*node-set*) bat denotatzen du
- Bide-espresioak:
 - ◆ Absolutuak (/ du hasieran)
 - ◆ Erlatiboak (ez du / hasieran)
 - ◆ Bi kasuetan tarteko barrek bidea osatzen duten kokapen-urratsak (*location step*) bereizten dituzte:

/urratsa/urratsa/...

urratsa/urratsa/...

Adabegiak adierazteko bideak (II)

- Kokapen-urratsak banan-banan ebaluatuko dira, ezkerretik eskuinera
- Urrats bakoitza **uneko adabegi-multzoaren** aurka ebaluatuko da
- **Uneko adabegi-multzoa**: kokapen-bidea...
 - ◆ ...absolutua bada: **adabegi erroa**
 - ◆ ...erlatiboa bada: **espresioa erabiltzen ari deneko adabegia**

Adabegiak adierazteko bideak (III)

■ **Kokapen-urratsa** (*location step*):

- ◆ Ardatza (*axis*): urratsak zehaztutako adabegien eta uneko adabegiaren arteko erlazioa
- ◆ Adabegi-egiaztapena (*node test*): urratsak zehaztutako adabegi(ar)en mota(k) eta deitura hedatua(k)
- ◆ Predikatuak (zero edo gehiago): urratsak zehaztutako adabegien multzoa areago murrizteko espresioak

■ Sintaxia:

ardatza::adabegi-egiaztapena[predikatua]

■ Adibidea:

child::price[.=9.90]

Ardatzak eta adabegi-egiaztapenak

- Ardatzak adabegi-multzo bat zehazten du uneko adabegiarekiko
- Adabegi bat ardatz baten arabera identifikatzeko adabegi-egiaztapena erabiltzen da. Adabegi-egiaztapen bat deituraz edo motaz egiten ahal da.
- Adabegien mota-egiaztapena:
 - ◆ **comment ()**: iruzkin motako adabegia
 - ◆ **node ()**: edozein adabegi, atributuak edo izen-espazioak izan ezik
 - ◆ **processing-instruction ()**: prozesatze-agindua
 - ◆ **text ()**: testu motako adabegia

Ardatzak

- `ancestor`: uneko adabegiaren arbaso guztiak (gurasoa, gurasoaren gurasoa...)

Oharra: adabegi erroa beti barne, non eta ez den bera unekoa

- `ancestor-or-self`: uneko adabegia gehi arbaso guztiak

- `attribute`: uneko adabegiaren atributu guztiak

- `child`: uneko adabegiaren ume guztiak

Oharra: ez dira umetzat hartzen atributu-adabegiak (ez eta izen-espazioko adabegiak ere)

Ardatzak (II)

- descendant: uneko adabegiaren ondorengo guztiak (umeak, umeen umeak...)

Oharra: inoiz ez ditu hartzen atributu-adabegiak

- descendant-or-self: uneko adabegia gehi ondorengo guztiak
- following: uneko adabegiaren atzetik (dokumentuan) datorren guztia
- following-sibling: uneko adabegiaren atzetik (dokumentuan) datozen senide guztiak

Oharra: uneko adabegia atributu-adabegia bada, ardatz hau hutsik egongo da

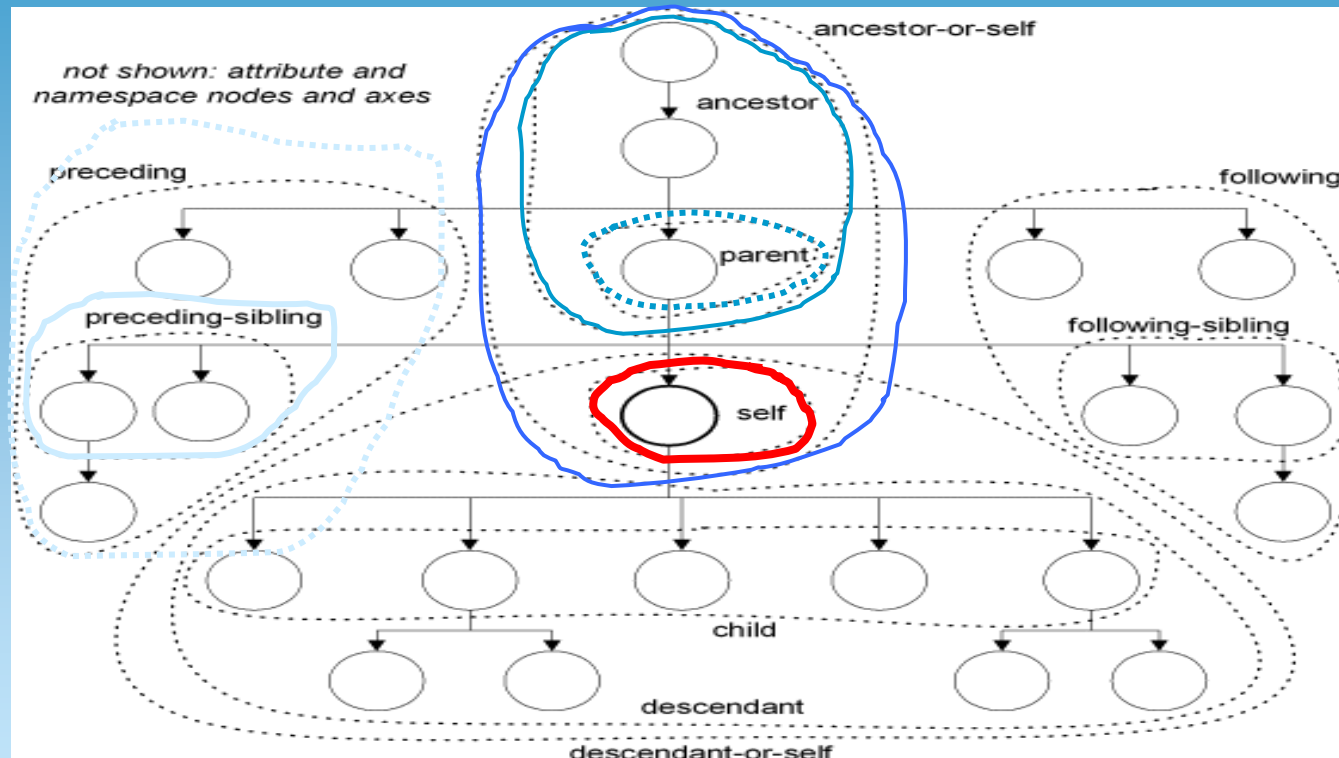
Ardatzak (III)

- `parent`: uneko adabegiaren gurasoa
- `preceding`: uneko adabegiaren aurretik (dokumentuan) datorren guztia
- `preceding-sibling`: uneko adabegiaren aurretik (dokumentuan) datozen senide guztiak

Oharra: uneko adabegia atributu-adabegia bada, ardatz hau hutsik egongo da

- `self`: uneko adabegia

Ardatzak, grafikoki



ancestor, descendant, following, preceding eta self ardatzen **bildura dokumentu osoa da** (atributuak eta izen-espazioak alde batera utziz), eta **beren arteko ebakidurak hutsak dira**

Adabegien arteko ordena ardatzetan

- Ardatz batzuek zuhaitzeko **adabegien artean ordena** bat dagoela (dokumentuaren araberrakoa) suposatzen dute :

zuhaitza aurreordenaz korritzen da, ezkerretik eskuinera (testu-adierazpidean aurkituko dugun ordena bera!)

- Ardatz batzuetan (following, child...) ordena **aurreranzkoa** da, baina beste batzuetan (ancestor, preceding-sibling...) **atzeranzkoa**; hau inportantea da posizioekin-eta jokatzeko denerako!

Adibideak

- **child::cd**

uneko adabegiaren umeen artean, cd elementu guztiak (ez bada horrelakorik, adabegi-multzo hutsa izango da emaitza)

- **attribute::src**

uneko adabegiaren src atributua (uneko adabegiak ez badu src atributurik, adabegi-multzo hutsa)

- **child::***

uneko adabegiaren ume guztiak

Adibideak (II)

- **attribute::***

uneko adabegiaren atributu oro

- **child::text()**

uneko adabegiaren testu-adabegi umeak

- **child::node()**

uneko adabegiaren ume guztiak

- **descendant::cd**

uneko adabegiaren ondorengoetan, cd
elementu guztiak

Adibideak (III)

- **ancestor::cd**

uneko adabegiaren arbasoetan, cd guztiak

- **ancestor-or-self::cd**

uneko adabegiaren cd arbasoak, gehi, unekoa cd bada, unekoa bera ere bai

- **child::* / child::price**

uneko adabegiaren price "biloba" (umearen ume) guztiak

- /

dokumentuaren adabegi erroa

Predikatuak: []

- Predikatu batek adabegi-multzo bat iragazi egiten du, adabegi-multzo berri bat sortuz
- `child::price[text()=9.90]`
unekoaren umeen artean, balioztat 9.90 duten `price` elementuak
- `child::cd[position()=1]`
unekoaren lehen `cd` umea
- `child::cd[position()=last()]`
unekoaren azken `cd` umea

Predikatuak (II)

- `child::cd[position()=last()-1]`
unekoaren azkenaurreko cd umea
- `child::cd[position()<6]`
unekoaren cd umeetan, lehenengo bostak
- `/descendant::cd[position()=7]`
dokumentuko zazpigarren cd elementua
- `child::cd[attribute::type="classic"]`
unekoaren cd umeetan, classic baliodun type atributu bat duten guztiak

Kokapen-bideak: sintaxi laburtua

- Kokapen-bideotan laburtzapenak erabil daitezke
- `child :: ardatza, esaterako, beti utz daiteke aipatu gabe kokapen-bide batean, besterik ezean child :: ulertuko baita`

`cd ≡ child :: cd`

Kokapen-bideak: sintaxi laburtua (II)

	<code>child::</code>	cd ≡ <code>child::cd</code>
@	<code>attribute::</code>	cd[@type="classic"] ≡ <code>child::cd</code> [<code>attribute::type="classic"</code>]
.	<code>self::node()</code>	./cd ≡ <code>self::node()/</code> <code>descendant-or-self::node()/</code> <code>child::cd</code>
..	<code>parent::node()</code>	../cd ≡ <code>parent::node()/child::cd</code>
//	<code>/descendant-or-self::node()</code>	//cd ≡ <code>/descendant-or-self::node()/</code> <code>child::cd</code>

Adibideak, labur

- **cd**

uneko adabegiaren `cd` elementu guztiak

- *****

uneko adabegiaren ume guztiak

- **text ()**

uneko adabegiaren ume diren testu-adabegi guztiak

- **@src**

uneko adabegiaren `src` atributua

- **@***

uneko adabegiaren atributu guztiak

- **cd[1]**

uneko adabegiaren lehen `cd` umea

Adibideak, labur (II)

- `cd[last()]`
uneko adabegiaren azken `cd` umea
- `*/cd`
uneko adabegiaren umeen ume diren `cd` guztiak
- `/book/chapter[3]/para[1]`
liburuaren 3. kapituluko lehen paragrafoa
- `//cd`
adabegi erroaren ondorengoetan, `cd` guztiak
- `.`
uneko adabegia
- `././cd`
uneko adabegiaren ondorengoetan, `cd` diren guztiak

Adibideak, labur (III)

- ..
uneko adabegiaren gurasoa
- ../@src
uneko adabegiaren gurasoaren src atributua
- cd[@type="classic"]
uneko adabegiaren cd umeetan, classic balioa duen type atributua dutenak
- cd[@type="classic"][5]
uneko adabegiaren cd umeen artean classic balioa duen type atributu bat dutenetan, bosgarrena
- cd[5][@type="classic"]
unekoaren bosgarren cd umea, baldin eta classic balioa duen type atributu bat badu
- cd[@type and @country]
unekoaren cd umeetan, type atributu bat eta country atributu bat duten guztiak

Espresioak

- Aritmetikoak: +, -, *, div, mod
- Berdin/desberdin: =, !=
- Erlazionalak: <, <=, >, >=
- Boolearrak: or, and

Espresioak (II)

- Adabegi-multzo batekiko konparazioa:
 - ◆ $=: true$, adabegi-multzoak adabegiren bat baldin badauka emandako balioa duena
 - ◆ $! =: true$, adabegi-multzoak adabegiren bat baldin badauka emandako balioa ez duena

Beraz, posible da adabegi-multzoa berdina eta desberdina izatea aldi berean!!!

Kontuz ibili, beraz, = eta != eragileekin, beti ez baita "espero" dena lortzen!!!

Funtzioak: adabegi-multzoak

- **count** (*adabegi-multzoa*) => *zenbakia (number)*
 - ◆ hautaturiko elementuen kopurua
- **id** (*balioa*) => *adabegi-multzoa (node-set)*
 - ◆ emandako identifikadorea daukan elementua (bakarra)
- **last** () => *zenbakia (number)*
 - ◆ testuinguruaren tamaina (espresioaren ebaluazio-testuinguruaren menpeko balioa)
- **local-name** (*[adabegia]*) => *string*
 - ◆ emandako (edo, argumenturik ezean, uneko) adabegiaren deitura hedatuaren parte lokala (bi puntuen ondokoa, *namespace*-ri dagokion aurrizkia, alegia)

Funtzioak: adabegi-multzoak (II)

- **name** (*[adabegia]*) => *string*
 - ◆ emandako adabegiaren izena (edo unekoarena, argumenturik ezean)
- **namespace-uri** (*[adabegia]*) => *string*
 - ◆ emandako (edo, argumenturik ezean, uneko) adabegiaren deitura hedatuaren *namespace URI* delakoa (*namespace* erazagupeneko, normalean)
- **position** () => *zenbakia (number)*
 - ◆ testuinguruaren posizioa (1etik hasita; espresioaren ebaluazio-testuinguruaren menpeko balioa)

Funtzioak: string-ak

- **concat**(*str1*, *str2*, *str3...*) => *string*
 - ◆ argumentuak kateatuz eratutako string-a
 - ◆ `concat('The', ' ', 'XML')` => `'The XML'`
- **contains**(*str1*, *str2*) => *boolean*
 - ◆ *true*, lehen string-ak bigarrena badauka bere barruan; *false*, bestela
 - ◆ `contains('XML', 'X')` => *true*
- **normalize-space**([*str1*]) => *string*
 - ◆ emandako string-etik (edo uneko adabegiaren string-balioetik, argumenturik ezean) aurreko eta atzeko zuriuneak kendu, eta zuriune bat baino gehiagoko segidak zuriune bakarreko bihurtzen ditu
 - ◆ `normalize-space(' The XML')` => `'The XML'`

Funtzioak: string-ak (II)

- **starts-with(*str1*, *str2*) => boolean**
 - ◆ *true*, lehen string-a bigarrenarekin hasten bada; *false*, bestela
 - ◆ `starts-with('XML', 'X') => true`
- **string([*balioa*]) => string**
 - ◆ emandako balioa string bihurtzen du (uneko adabegia hartzen du besterik ezeko argumentutzat)
 - ◆ `string(3.14) => '3.14'`
- **string-length([*str*]) => string**
 - ◆ *str*-ren karaktere kopurua (uneko adabegiaren string-balioa hartzen da argumentutzat besterik ezean)
 - ◆ `string-length('Beatles') => 7`

Funtzioak: string-ak (III)

- **substring(*str*, *nondik*, [*kop*]) => string**
 - ◆ *nondik*-etik hasita *kop* karaktereko azpikatea (edo bukaerarainokoa, hirugarren argumenturik ezean)
 - ◆ `substring('Beatles', 2, 4) => 'eatl'`
- **substring-after(*str1*, *str2*) => string**
 - ◆ *str2* *str1*-en estreina agertzen denetik bukaerarainoko azpikatea (*str1*-ekoa)
 - ◆ `substring-after('12/10', '/') => '10'`
- **substring-before(*str1*, *str2*) => string**
 - ◆ *str2* *str1*-en estreina agertzen dena baino lehenagoko azpikatea (*str1*-ekoa)
 - ◆ `substring-before('12/10', '/') => '12'`
- **translate(*str1*, *str2*, *str3*) => string**
 - ◆ *str2*-n ageri diren *str1*-eko karaktereak ordezkatzeko dituzten *str3*-koekin
 - ◆ `translate('12:30', ':', '!') => '12!30'`

Funtzioak: zenbakiak

- **ceiling(*zenb*) => zenbakia (*number*)**

- ◆ *zenb* baino handiagoa (edo berdina) den osorik txikiena
- ◆ `ceiling(3.14) => 4`

- **floor() => zenbakia (*number*)**

- ◆ *zenb* baino txikiagoa (edo berdina) den osorik handiena
- ◆ `floor(3.14) => 3`

- **number([*balioa*]) => zenbakia (*number*)**

- ◆ emandako balioa (edo uneko adabegiaren string-balioa, argumenturik ezean) zenbaki bihurtzen du
- ◆ `number(price)`

Funtzioak: zenbakiak (II)

- **round**(*balioa*) => *zenbakia (number)*
 - ◆ *balioa*-tik gertuen dagoen osoa
 - ◆ `round(3.14) => 3`
- **sum**(*adabegi-multzoa*) => *zenbakia (number)*
 - ◆ emandako adabegi-multzoko nodo guztien string-balioen batura (string-balio horiek zenbaki bihurturik, noski)
 - ◆ `sum(//cd/price)`

Funtzioak: boolearrak

- **boolean(*balioa*) => *boolean***

- ◆ *balioa* boolear bihurtzen du:

- ◆ zenbakiak: 0 => *false*

- ◆ string-ak: hutsa => *false*

- ◆ boolearrak: ez dira aldatzen

- ◆ adabegi-multzoak: hutsa => *false*

- **false() => *boolean* (=false)**

- ◆ *false* itzultzen du (konstante booleartzat erabil daiteke)

- ◆ `number(false()) => 0`

- **not(*baldintza*) => *boolean***

- ◆ *true*, baldintza *false* bada; *false*, bestela

- ◆ `not(false())`

Funtzioak: boolearrak (II)

- **true()** => *boolean (=true)*
 - ◆ *true* itzultzen du (konstante booleartzat erabil daiteke)
 - ◆ `number(true()) => 1`
- **lang(*h*)** => *boolean*
 - ◆ *true*, uneko adabegiaren hizkuntza *h* edo *h*-ren "azpihizkuntza" bat baldin bada; *false*, bestela (baita uneko adabegiak hizkuntza aitorturik ez badauka ere)
 - ◆ `xml : lang` atributuan oinarritzen da funtzio hau:
 - ◆ `eu, es, en, en-US, fr-CA...`

Ohar pare bat

- **Mota-bihurketa automatikoak:** funtzio hauetako askotan, argumentuak string edo zenbaki izan behar duelarik ez bada, `string()` edo `number()` funtzioa (hurrenez hurren) aplikatuko zaio, eta bihurtu egingo da funtzioa aplikatu aurretik (eta gauza bera boolearrekin).
- **Adabegi baten string-balioa:** Unicode karaktereen sekuentzia.
 - ◆ testu-adabegia: bertako testua osatzen duten karaktereak
 - ◆ atributu-adabegia: atributuaren balioa
 - ◆ iruzkina: bertako testua (mugatzailerik gabe)
 - ◆ elementua: umeen testu-balioen kateadura (dokumentu-adabegiaren kasuan, dokumentuko testu osoa, marka guztiak kenduta)

XLink, XPointer eta XQuery

XPath-ekin harreman estua
duten hiru lengoiaia

XLink, XPointer eta XPath

Hiru lengoia, hiru geruza:

- **XLink:** dokumentu barruko eta dokumentu arteko estekak orokortzeko lengoia
- **XPointer:** XPath-en hedapena, XLink-en erabilia urruneko baliabideen kokapena zehazteko
- **XPath:** XML zuhaitzetako adabegi eta zatiak zehazteko lengoia (oinarrizkoa)
- Teknologia hauek guztiak estandarrak dira (edo estandar izateko bidean daude) dagoeneko, baina ez daude oraindik erabat garatuta eta zabaldua, eta ezta tresna askotan inplementatuta ere (XPath izan ezik)

XLink

- HTMLko estekaren generalizazio bat
- Abstrakzio-maila altuagoa (XML orokorra helburu, ez esteka soilak)
- Ahalmen espresibo handiagoa (helburu bat baino gehiago, portaera bereziak, estekategiak...)
- XPointer erabiltzen du baliabideak erreferentziatzeko

XLink: adibidea

```
<mylink xmlns:xlink="http://www.w3.org/1999/xlink"
  xlink:type="extended">
  <myresource xlink:type="locator"
    xlink:href="students.xml#Fred" xlink:label="student"/>
  <myresource xlink:type="locator"
    xlink:href="teachers.xml#Joe" xlink:label="teacher"/>
  <myarc xlink:type="arc"
    xlink:from="student" xlink:to="teacher"/>
</mylink>
```

- Urruneko bi baliabide lotzen dituen **esteka "hedatua"** (**extended**)
- **xlink** aurrizkia (izen-espazioa) erabiltzen da, normalean, XLink prozesatzaileari estekak-eta zein diren erakusteko
- **locator**: urruneko baliabideak zehazteko (XPointer espresioa)
- **arc**: nondik nora doan esteka

XPointer

- XPath-en hedapena, estekatzeko helburuetarako egokia
- XPath espresioen eta URIen arteko harremana zehazten du
- Helbideratze erlatiboa: aingurarik gabeko lekuak erakusteko/lotzeko bidea eskaintzen du
- Malgua eta sendoa: helburu-dokumentuan aldaketarik egonda ere, XPointer/XPath espresioek maiz zuzen irauten dute
- Testu-adabegietako azpikateak nahiz zuhaitz-zati osoak erakusteko ahalmena du

XPointer: adibidea

URI

`http://www.foo.org/bar.xml#xpointer(article/section[position()<=5])`

**XPointer espresioa
(XPath!)**

XPointer zati-identifikatzailea

XPointer espresioak
article erroko
elementuko lehen bost
atalak (section) erakusten
ditu

XPath: oinarria

- zuhaitz-adabegi eta -zatiak zehazteko lengoia deklaratioa
- erabiltzen da...
 - ◆ XPointer-en (helbideak adierazteko)
 - ◆ XSLn (patroi-parekatzean)
 - ◆ XML Schema-n (esparruak zehazteko eta unibokotasuna lortzeko)
 - ◆ XQuery-n (hautatze eta iteraziorako)
 - ◆ XLink-en

XQuery vs. XPointer/XPath

- XQuery-k XPointer eta XPath gogorarazten ditu, baina beste helburu batzuk ditu:
 - ◆ XQuery: SQL modukoa, XML dokumentu edo datu-baseen galdeketarako
 - ◆ XPointer/XPath: informazio ezagunean barrena egin beharreko helbideratzea dute helburu

XQuery: adibidea (*FLWOR*)

```
for $d in document("depts.xml")//deptno
let $e := document("emps.xml")//employee[deptno = $d]
Where count($e) >= 10
Order by avg($e/salary) descending
Return
  <big-dept>
    {
      $d,
      <headcount>{count($e)}</headcount>,
      <avgsal>{avg($e/salary)}</avgsal>
    }
  </big-dept>
```

- **Galdera: lortu 10 langile baino gehiago dituzten sailen zerrenda, batez besteko soldaten arabera ordenatua**
- for: //deptno sail-zenbakien lista sortu eta \$d aldagaiari lotu
- let: lotu \$e aldagaian \$d sailekoak diren employee elementuak => (\$d, \$e) lotura-bikoteen zerrenda zehaztu da honela
- where, order: SQLkoen antzekoak
- return: emaitzaren eraikuntza (XML) => sail "handiak", zenbat langile dituzten eta batez besteko soldata zehaztuz

UEU. Iruñea, 2004ko uztaila.

UPV/EHU. Donostiako Informatika Fakultatea. Softwarea garatzeko tresna aurreratuak (2004/07)

XPath espresioak: lehen ariketa

- Kargatu, Oxygen-en bidez, *g_.xml* dokumentua
- Adibidea: dokumentuko sarrera guztiak (`entry`) hautatu:
`//entry`
- Idatzi eta probatu XPath galdera hauek:
 1. Kategoriatzat (`pos`) "sf." duten artikuluak (`entry`) hautatu
 2. Zenbat dira goiko espresioak hautatu dituenak? Idatzi hori kalkulatu duen XPath espresioa
 3. Hautatu "tiene " edo "Tiene " beren barruan duten definizio guztiak
 4. Hautatu dokumentuko artikuluetan ageri diren pluraleko forma (`orth`) guztiak (`form type="infl"` eta `<number>(pl.)</number>`)
 5. Aditzen artean (kategoria "v" letrarekin hasten da), hautatu eredutzat (`itype`) "(4)" dutenak. Lortu emaitzatzat `orth` elementuaren balioa.

XPath espresioak: bigarren ariketa

- Kargatu orain *hiztegia.xml* dokumentua.
- Idatzi eta probatu XPath galdera hauek:
 1. "ad." kategoriako (`pos`) artikuluak (`entry`).
 2. Zenbat definiziotan (`def`) ageri den "ahur" hitza.
 3. Zenbat definizio (`def`) edo adibidetan (`q` elementua eg barruan) ageri den "ardo" hitza.
 4. Zenbat artikulu (`entry`) dauden kategoriatzat (`pos`) "adlag." dutenak.
 5. "a0002" identifikadorea (`id` atributua, `entry` elementuan) duen sarreraren kategoria (`pos`).

XPath espresioak: bigarren ariketa (II)

6. Hiru adiera (`sense`) edo gehiago dituzten artikuluen (`entry`) sarrerak (`orth`).
7. Sinonimoren bat (`xr/ref`, `syn` klasekoa) duten sarrerak (`orth`).
8. Zein sarreraren (`orth`) definizio (`def`) edo adibideetan (`q` elementua `eg` barruan) ageri den "ardo" hitza.
9. "ahulezia" hitza sinonimotzat (`xr/ref`, `syn` klasekoa) duen sarrera (`orth`).

XPath: zein espresioek hautatzen ditu gorritz dauden adabegiak?

```
<AAA>
  <BBB>
    <CCC />
    <ZZZ />
  </BBB>
  <XXX>
    <DDD>
      <EEE />
      <FFF>
        <HHH />
        <GGG>
          <JJJ>
            <QQQ />
          </JJJ>
          <JJJ />
        </GGG>
        <HHH />
      </FFF>
    </DDD>
  </XXX>
  <CCC>
    <DDD />
  </CCC>
</AAA>
```

```
<AAA>
  <BBB>
    <CCC />
    <ZZZ />
  </BBB>
  <XXX>
    <DDD>
      <EEE />
      <FFF>
        <HHH />
        <GGG>
          <JJJ>
            <QQQ />
          </JJJ>
          <JJJ />
        </GGG>
        <HHH />
      </FFF>
    </DDD>
  </XXX>
  <CCC>
    <DDD />
  </CCC>
</AAA>
```

XPath: zein espresioak hautatzen ditu gorritz dauden adabegiak? (II)

```
<AAA>
  <BBB>
    <CCC/>
    <ZZZ>
      <DDD/>
    </ZZZ>
  </BBB>
  <XXX>
    <DDD>
      <EEE/>
      <DDD/>
      <CCC/>
      <FFF/>
      <FFF>
        <GGG/>
      </FFF>
    </DDD>
  </XXX>
  <CCC>
    <DDD/>
  </CCC>
</AAA>
```

```
<AAA>
  <BBB>
    <CCC/>
    <DDD/>
  </BBB>
  <XXX>
    <DDD>
      <EEE/>
      <DDD/>
      <CCC/>
      <FFF/>
      <FFF>
        <GGG/>
      </FFF>
    </DDD>
  </XXX>
  <CCC>
    <DDD/>
  </CCC>
</AAA>
```