



UNIUNEA EUROPEANĂ



GUVERNUL ROMÂNIEI



Instrumente Structurale  
2007-2013



## Platformă de e-learning și curriculum e-content pentru învățământul superior tehnic

Proiect nr. 154/323 cod SMIS – 4428 cofinanțat de prin Fondul European de Dezvoltare Regională “Investiții pentru viitorul dumneavoastră”.

**Programul Operațional Sectorial Creșterea Competitivității Economice - POS CCE**

---



UNIUNEA EUROPEANĂ



GUVERNUL ROMÂNIEI



Instrumente Structurale  
2007-2013



Platformă de e-learning și curriculum e-content  
pentru învățământul superior tehnic

Interacțiunea om-calculator

8. XPath

---

# Adresarea către interiorul documentelor XML

- Limbajul XPath
- Limbajul XPointer

# Limbajul XPath

XPath consideră un document XML ca un arbore format din șapte tipuri de noduri:

- **Rădăcina** - copii: nodul element corespunzător documentului și, eventual noduri pentru instrucțiuni de procesare și noduri comentariu.
- **Element**. Pentru fiecare element din document există câte un nod element. Copii - alte noduri element, noduri pentru instrucțiuni de procesare, comentarii și noduri pentru text. Fiecare nod element poate avea un identificator unic, dacă este specificat prin tipul ID.

## Limbajul XPath (cont.)

- **Atribut**. Fiecare nod element are o mulțime de noduri atribut, corespunzătoare atributelor sale. Aceste noduri atribut au ca părinte nodul element respectiv dar ele nu sunt copii ai elementului părinte.
- **Spațiu de nume**.
- **Instrucțiuni de prelucrare**
- **Comentariu**
- **Text** - secvențe maximale de date de tip caracter.

## Limbajul XPath (cont.)

Principala construcție XPath este expresia. În urma evaluării expresiilor XPath, se obțin obiecte de tipurile:

- un nod;
- mulțime de noduri;
- boolean;
- număr (în virgulă mobilă);
- șir de caractere.

## Limbajul XPath (cont.)

Evaluarea fiecărei expresii XPath se face într-un context format din:

- un nod context;
- o pereche de numere întregi care reprezintă o poziție și o mărime a contextului, prima fiind mai mică decât cea de-a doua;
- legări de variabile;
- o bibliotecă de funcții;
- mulțimea declarațiilor de spații de nume.

## Limbajul XPath (cont.)

Principalul tip de expresie XPath este **calea către o locație** (“location path”). Ea include convențiile uzuale folosite pentru specificarea căilor către fișiere în sistemele de operare.

Căile către locații sunt o secvență de pași separați de caracterul “/”. Ele pot fi:

- Relative - căi care încep de la nodul context curent.
- Absolute - căi care pleacă din rădăcina documentului, fapt indicat prin faptul că încep cu “/”.



## Limbajul XPath (cont.)

***axa::nod[predicat\*]***

Fiecare pas are trei părți:

- o axă,
- un tip de nod
- zero sau mai multe predicate.

## Limbajul XPath (cont.)

Axa specifică direcția în care se face pasul curent. Ea poate fi:

- child, nodul copil, aceasta fiind axa implicită;
- descendant, adică copii sau copiii copiilor, pe oricâte nivele;
- parent
- ancestor, adică părinte sau părintele părintelui ș.a.m.d.;
- following-sibling
- preceding-sibling
- following
- preceding
- attribute, adică atributele nodului context;
- namespace
- self
- descendent-or-self
- ancestor-or-self

## Limbajul XPath (cont.)

Expresiile XPath pot fi folosite diverse funcții predefinite.

- `number last()` - ultimul element dintr-o mulțime
- `number position()` - poziția unui element
- `number count(node-set)` - numărul de elemente
- `node-set id(object)` - nodul cu identificatorul dat ca argument
- `string name(node-set?)`
- `string string(object?)`
- `boolean starts-with(string, string)`
- `boolean contains(string, string)`
- `string substring-before(string, string)`
- `number string-length(string?)`

## Limbajul XPath (cont.)

Abrevieri	Ce se consideră față de nodul context
*	toți copii
text()	copiii de tip text
@nume	atributul denumit "nume"
@*	toate attributele
nume[i]	al i-lea copil "nume"
nume[last()]	ultimul copil "nume"
*/nume	nepoții "nume"
//nume	descendenții "nume"
.	nodul context
..	Părintele
../@nume	atributul "nume" al părintelui
nume[@atr="val"]	copii "nume" care au atributul "atr" cu valoarea "val"

```

<?xml version="1.0"?>

<bib>
<webdoc src="http://concept.cs.uah.edu/CG/cg-standard.html" abr="CG">
Conceptual Graphs </webdoc>
<webdoc abr="CYC" src="http://www.cyc.org"> CYC</webdoc>
<webdoc abr="DOM" src="http://www.w3.org/DOM"> DOM </webdoc>
<webdoc abr="KIF" src="http://logic.stanford.edu/kif/kif.html">
Knowledge Interchange Format </webdoc>
<webdoc abr="KQML" src="http://www.cs.umbc.edu/kqml"> KQML </webdoc>
<webdoc abr="RDF" src="http://www.w3.org/RDF"> RDF Specification
<an>1999</an></webdoc>
<webdoc abr="RDFS" src="http://www.w3.org/TR/WD-rdf-schema"> RDF Schema
Specification</webdoc>
<articol abr="Sho93">
  <autor> Y. Shoham, Agent-oriented programming </autor>
  <revista nr="60">
    <titlu>Artificial Intelligence</titlu>
    <an>1993</an>
    <pp> 51-92</pp>
  </revista>
</articol>
<webdoc abr="XML" src="http://www.w3.org/XML"> XML </webdoc>
<carte abr="XMLc" isbn="1-861003-4-12">
  <autor> Cagle, K. </autor><autor>Gibbons, D.</autor><autor> Hunter,
D.</autor><autor> Ozu, N.</autor><autor>Pinnock, J.
  </autor><autor>Spencer, P.</autor>
  <titlu> Beginning XML </titlu>
  <editura>Wrox Press,</editura>
  <an> 2000</an>
</carte>
</bib>

```

Expresie XPath	Rezultat obținut
/bib/webdoc	Conceptual Graphs
//an	1999
/bib/articol/*	Y. Shoham, Agent-oriented programming
/bib/webdoc/following-sibling::webdoc	CYC
/bib/webdoc/attribute::*	<a href="http://concept.cs.uah.edu/CG/cg-standard.html">http://concept.cs.uah.edu/CG/cg-standard.html</a>
//an/../@abr	RDF
/bib/webdoc[6]	RDF Specification 1999
/bib/webdoc[4]/following-sibling::*	KQML
/bib/webdoc[@*]	Conceptual Graphs
//*[@nr]	Artificial Intelligence 1993 51-92
/bib/webdoc[6][@abr]	RDF Specification 1999
/bib/webdoc/text()	Conceptual Graphs
/bib/webdoc[last()]	XML
//*[@count(*)>2]/*[2]	CYC

## Limbajul XPointer

În HTML se poate face o referire (de exemplu, o legătură) doar la un întreg document sau la un punct specificat prin #ancora (și definit în document prin <a name="ancora">).

Expresiile XPointer sunt extensii ale XPath și pot fi adăugate la un URI, după caracterul "#", ca în următorul exemplu:

**http:// . . . /b1.xml#xpointer(/bib/articol)**

## Limbajul XPointer (cont.)

- XPointer este un limbaj care extinde XPath în sensul că permite adresarea nu numai a unui nod sau a unei mulțimi de noduri (ca în XPath) ci și a unui punct sau a unui fragment dintr-un nod al unui document.
- Un **punct** dintr-un document este precizat printr-o pereche formată dintr-un nod container și un index. Dacă nodul container are fii, indexul se referă la al câtelea fiu este considerat. Dacă nodul container nu are fii, de exemplu, este un nod text, indexul este considerat față de șirul de caractere care formează nodul.



## Conexiuni între documente. Limbajul XLink

Specificarea legăturilor în XLink se face folosind documente XML bine formate, pe baza unui repertoriu de attribute, după cum urmează:

- “type” este un atribut care specifică **tipul** legăturii. El poate avea ca valoare:
  - “simple”, pentru legăturile simple, între două resurse.
  - “extended”, pentru legături extinse.

## Legaturi extinse

Pentru acest tip de legături se pot defini elemente pentru specificări suplimentare, conform următoarelor patru atribute:

- “locator”, pentru resurse externe,
- “arc”, pentru specificarea direcțiilor de traversare a legăturilor,
- “resource”, pentru specificarea de resurse locale,
- “title”, pentru a specifica un element pentru titlu (în locul atributului cu același nume, de care se va vorbi un pic mai jos).

# Atribute referitoare la semantica legăturii

- “role”,
- “title”.

## Atribute referitoare la comportarea legăturii

- “actuate”, precizează când va căutată noua resursă. Poate avea ca valoare:
  - “onLoad”, similar cu “<img>” din HTML,
  - “onRequest”, similar cu “<a href=. . .>”,
  - “undefined” - specific aplicației.
- “show”, precizează cum va fi afișată noua resursă în urma parcurgerii unei legături. Poate avea ca valoare modalitatea de afișare:
  - “new”, într-o fereastră separată;
  - “replace”, prin înlocuire, ca la “<a href=. . .>” din HTML;
  - “embed”, ca la “<img>” din HTML;
  - “undefined”

# Atribute care descriu direcționalitatea

- “from”,
- “to”.