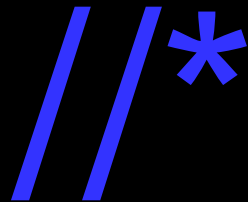


Tehnologii Web

un model de date pentru Web (II)



XML Infoset, acces via XPath, validări DTD

“Înțeleptul te învață prin faptele sale,
nu prin cuvintele sale.”

Chuang-Tzu

Cum modelăm (reprezentăm) datele?

xml-infoset

Scop:
specificarea unui model de date (abstract)
pentru XML

recomandare W3C, 2004

www.w3.org/TR/xml-infoset

xml-infoset

Se oferă un punct de vedere comun referitor la:

serializarea datelor semi-structurate

reprezentare internă (în memorie) \leftrightarrow fișier text XML

xml-infoset

Se oferă un punct de vedere comun referitor la:

crearea/folosirea de API-uri de procesare XML

...într-o manieră standardizată

xml-infoset

Se oferă un punct de vedere comun referitor la:

definirea unor specificații de nivel (mai) înalt
exemple: XPath, XSLT sau XML Schema

...care recurg la același model abstract XML

xml-infoset

Modelul asigura inter-operabilitatea diferitelor tehnologii, interfețe de programare (API-uri) și aplicații XML

xml-infoset: concepte

Document (*document information item*)

considerat a fi un arbore, cu rădăcina dată
de proprietatea **[document element]**

xml-infoset: concepte

Document (*document information item*)

are proprietatea **[children]** desemnând
lista de „lucruri” (*items*) din document,
în ordinea dată de algoritmul DFS

xml-infoset: concepte

Element

specifică un element XML

posedă proprietatea **[parent]** oferind informații despre elementul părinte căruia îi aparține

are asociată proprietatea **[children]**

xml-infoset: concepte

Element

proprietatea **[local name]** desemnează numele local al elementului aparținând unui spațiu de nume indicat via **[namespace name]**

prefixul spațiului de nume utilizat este stocat de proprietatea **[prefix]**

xml-infoset: concepte

Element

proprietatea **[attributes]** oferă lista neordonată a atributelor atașate

proprietatea **[namespace attributes]** specifică lista neordonată a atributelor **xmlns** asociate

xml-infoset: concepte

Atribut (*attribute*)

desemnează conceptul de atribut XML

numele și spațiul de nume atașat sunt specificate de proprietățile **[local name]** și **[namespace name]**, respectiv

xml-infoset: concepte

Atribut (*attribute*)

elementul căruia îi aparține este indicat
de proprietatea **[owner element]**

valoarea propriu-zisă a atributului
este specificată de **[normalized value]**

xml-infoset: concepte

Caractere (*characters*)

corespund informațiilor textuale
ale conținuturilor elementelor XML

proprietatea **[parent]** indică elementul căruia îi aparțin

proprietatea **[children]** conține datele-caracter propriu-zise

xml-infoset: concepte

Caractere (*characters*)

setul de caractere utilizat este desemnat de proprietatea **[character code]**

modul de procesare a spațiilor albe este specificat de proprietatea booleană **[element content whitespace]**

Document

```
• html "http://www.w3.org/1999/xhtml"
  ▼ • head
    • title
    • link "stylesheet"
    • meta "content-type"
    • meta "description"
    • meta "keywords"
    • meta "author"
  ▼ • body
    ▼ • article "top"
      ▼ • header
        ▼ • a "index.html"
          • img "web-logo.png"
          • h1
        ▼ • section "main"
          ▼ • aside "menu"
            ▼ • nav
              • a "web-film.html"
              • a "web-eval.html"
              • a "web-contact.html"
```

Element

Characters

xml-infoset: concepte

Instrucțiuni de procesare
(*processing instruction information item*)

Comentarii
(*comment information item*)

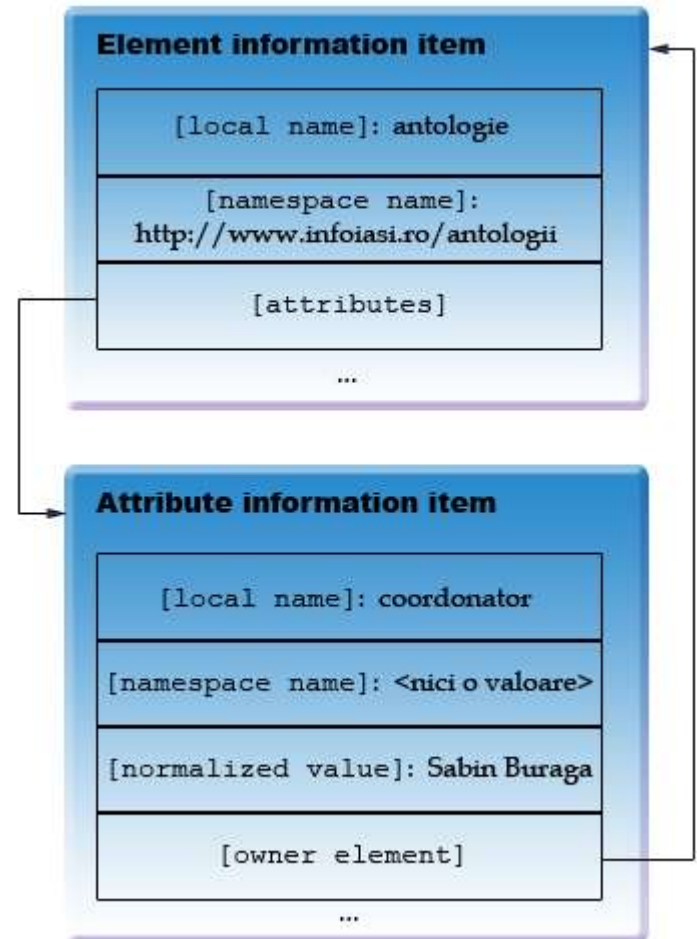
Declarația tipului de document
(*document type declaration item*)

xml-infoset: exemplu

Construcției

```
<antologie xmlns=  
  "http://www.infoiasi.ro/antologii"  
  coordinator="Sabin Buraga" />
```

îi corespunde *infoset*-ul:



În ce mod pot fi accesate datele stocate
de un document XML?

xsl: xpath

XPath 1.0

recomandare a Consorțiului Web (1999)

www.w3.org/TR/xpath

xsl: xpath

XPath 1.0

adreasează părți dintr-un document XML

oferă facilități de bază pentru manipularea datelor
(șiruri, numere,...)

operează la nivelul structurii abstracte
a documentelor XML (arborele)

xsl: xpath

expresia XPath – construcție de bază

evaluarea se realizează în funcție de context:

un nod al documentului XML

poziție

o funcție (predefinită sau specificată de programator)

o declarație a unui spațiu de nume

xsl: xpath

expresia XPath – construcție de bază

în urma evaluării expresiei, e returnat un obiect:

- o mulțime de noduri (*node-set*)

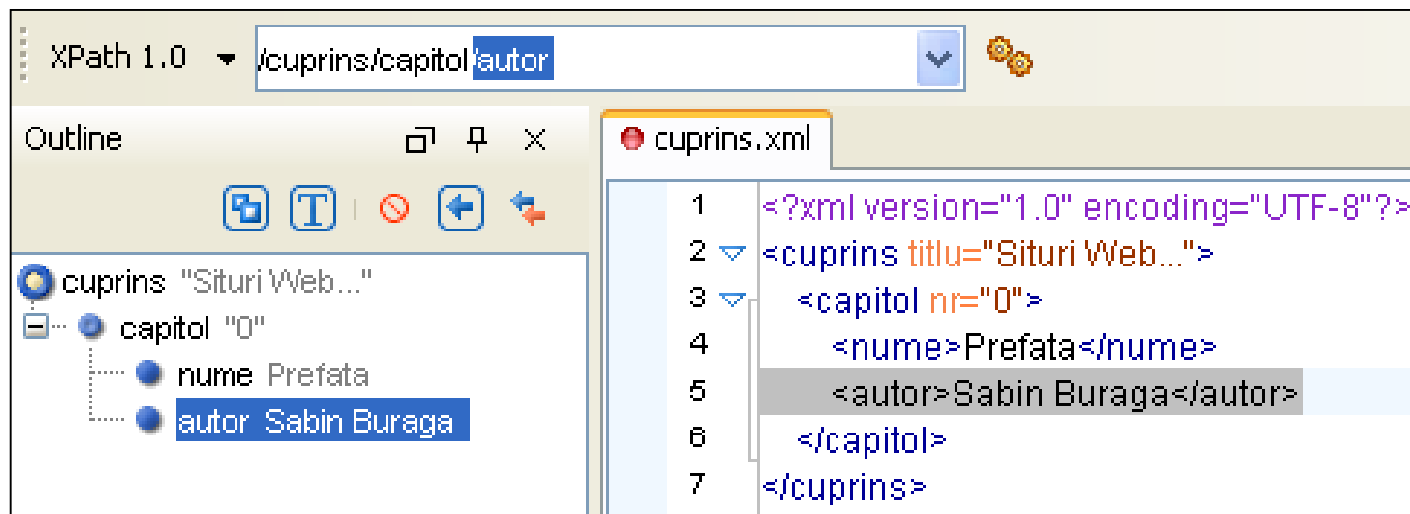
- o valoare logică – *true/false*

- un număr (*float*)

- un șir de caractere

xsl: xpath – operatori

Descendent /
/html/body/article
/cuprins/capitol/autor



xsl: xpath – operatori

Traversare recursivă //

//capitol

//cuprins

//div/a

xsl: xpath – operatori

Traversare recursivă //

//capitol

//cuprins

//div/a

Atenție la problemele de performanță!

De ce?

xsl: xpath – operatori

*Wildcard **

**/cuprins/*/nume
capitol/***

[5 \(19—23 martie\)](#)
[4 \(12—16 martie\)](#)
[3 \(5—9 martie\)](#)
[2 \(27 februarie—2 martie\)](#)
[1 \(20—24 februarie\)](#)

Calendar

Săptămâna 1

Supliment: *HTML, CSS, and*

- **Prezentare generală**

detalii organizat

- Curs: **Spațiul World Wide Web – aspecte a**
terminologie, model client/server, hipertext, U

- Laborator: De la limbajul (X)HTML către HTML,

- Instrumente: **XAMPP**, **Portable Apps**, **Sublin**
HTML Validator

Săptămâna 2

- Curs: **Programare Web – Protocolul HTTP.**
conținut dinamic prin CGI

HTTP, câmpuri-antet, mesaj, MIME, pr

The screenshot shows the FirePath extension interface in Firefox. At the top, there are tabs for 'Cons...', 'HTML', 'CSS', 'Script', 'DOM', 'Net', and 'FirePath'. The 'FirePath' tab is active. Below the tabs, there is a 'Top Window' dropdown, a 'Highlight' button, and an 'XPath' input field containing the query `/html/body/*/section`. To the right of the input field is an 'Eval' button. The main area displays the DOM tree with the following structure:

```
<document>
  <html lang="ro" xml:lang="ro" xmlns="http://www.w3.org/1999/xhtml">
    <head>
    <body>
      <article id="top" class="content">
        <header>
        <section class="main">
        <section id="notes">
        <footer class="vevent summary">
      </article>
    </body>
  </html>
</document>
```

The `<section id="notes">` node is highlighted in blue. Below the DOM tree, a large white box with a blue border contains the XPath query `/html/body/*/section`. At the bottom of the interface, a status bar shows a green checkmark and the text '2 matching nodes'.

evaluarea de expresii XPath
via extensia **FirePath** pentru Firefox

xsl: xpath – operatori

Nodul curent .

./capitol

xsl: xpath – operatori

Atribut @

capitol/@nr
table/@*

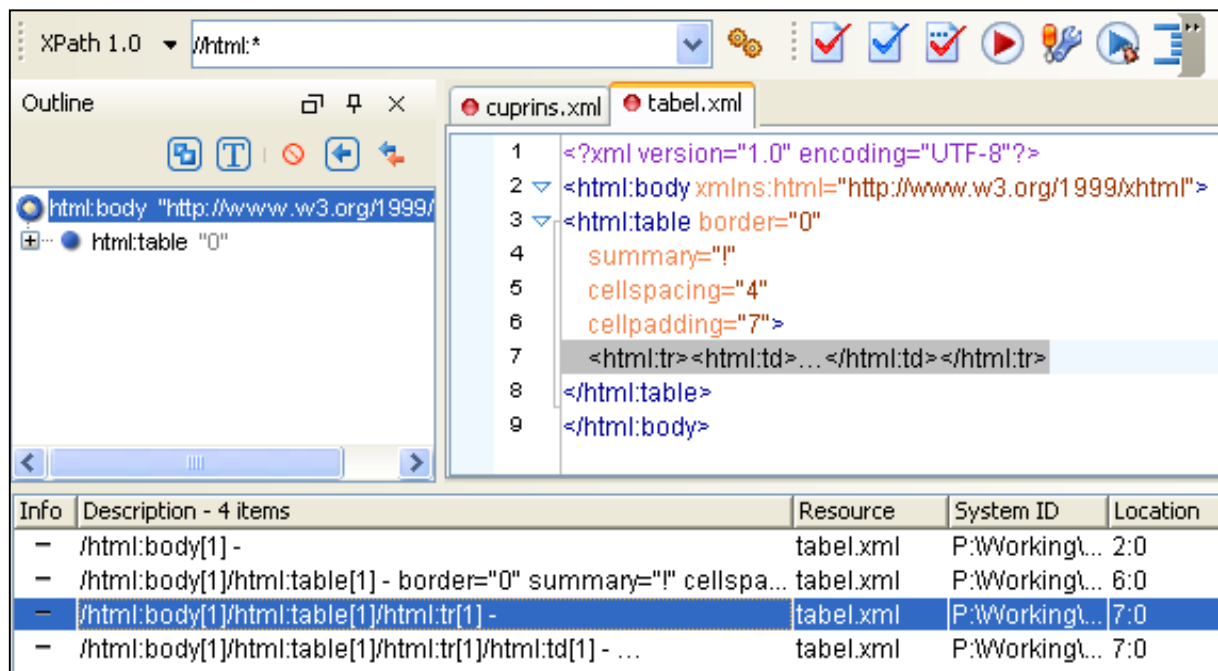
xsl: xpath – operatori

Spațiu de nume :

xsl:*

xsl: xpath – operatori

Spațiu de nume :
//html:*



The screenshot shows an XML editor interface. At the top, the XPath 1.0 query `//html:*` is entered. The main editor displays the XML content of `tabel.xml`, which includes an XML declaration, an `html:body` element, and an `html:table` element with a single `html:tr` containing an `html:td`. The `html:tr` element is highlighted. On the left, the Outline pane shows the document structure: `html:body` and `html:table`. At the bottom, the Info pane displays a table of 4 items representing the XPath results.

Info	Description - 4 items	Resource	System ID	Location
-	/html:body[1] -	tabel.xml	P:\Working\...	2:0
-	/html:body[1]/html:table[1] - border="0" summary="!" cellspa...	tabel.xml	P:\Working\...	6:0
-	/html:body[1]/html:table[1]/html:tr[1] -	tabel.xml	P:\Working\...	7:0
-	/html:body[1]/html:table[1]/html:tr[1]/html:td[1] - ...	tabel.xml	P:\Working\...	7:0

xsl: xpath – operatori

Filtru/index []

/cuprins/capitol[2]
//li[@class = "presentation"]

forma generală: [*expresie*]

5 (19–23 martie)
4 (12–16 martie)
3 (5–9 martie)
2 (27 februarie–2 martie)
1 (20–24 februarie)

Calendar

Săptămâna 1

Supliment: *HTML, CSS, and*

- **Prezentare generală**

detalii organizat

- Curs: **Spațiul World Wide Web – aspecte a**
terminologie, model client/server, hipertext, U
- Laborator: De la limbajul (X)HTML către HTML,
- Instrumente: **XAMPP**, **Portable Apps**, **Sublin**
HTML Validator

Săptămâna 2

- Curs: **Programare Web – Protocolul HTTP**,
conținut dinamic prin CGI

HTTP, câmpuri-antet, mesaj, MIME, pr

Top Window Highlight XPath: //ul[1] Eval

```
<document>
  <html lang="ro" xml:lang="ro" xmlns="http://www.w3.org/1999/xhtml">
    <head>
    <body>
      <article id="top" class="content">
        <header>
          <section class="main">
            <aside class="menu">
              <h2> Desfășurare </h2>
              <h3> Structură </h3>
              <p> Resursele disponibile mai jos completează și nu substituie laboratoare. </p>
            <p>
            <nav>
              <ul class="list">
            </nav>
            <h3> Calendar </h3>
            <aside class="menu supplement">
              <section id="week1">
                <h4> Săptămâna 1 </h4>
                <div class="info">
                  <ul>
            </div>
              </section>
            </aside>
          </section>
        </header>
      </article>
    </body>
  </html>
```

7 matching nodes

7 noduri (de tip Element) care satisfac expresia **//ul[1]**

xsl: xpath – operatori

Pentru valori logice și numerice pot fi folosiți
operatorii uzuali:

or and = != <= < >= >
+ - * div mod

comparațiile de șiruri sunt *case sensitive*

5 (19—23 martie)

4 (12—16 martie)

3 (5—9 martie)

2 (27 februarie—2 martie)

1 (20—24 februarie)

Calendar

Săptămâna 1

Supliment: *HTML, CSS, and c*

• Prezentare generală

detalii organizat

• Curs: Spațiul World Wide Web – aspecte ai

terminologie, model client/server, hipertext, UI

• Laborator: De la limbajul (X)HTML către HTML,

• Instrumente: XAMPP, Portable Apps, Sublin HTML Validator

Săptămâna 2

• Curs: Programare Web – Protocolul HTTP. conținut dinamic prin CGI

HTTP, câmpuri-antet, mesaj, MIME, pr

Top Window Highlight XPath: //li[@class="presentation"] Eval

```
<document>
  <html lang="ro" xml:lang="ro" xmlns="http://www.w3.org/1999/xhtml">
    <head>
    <body>
      <article id="top" class="content">
        <header>
        <section class="main">
          <aside class="menu">
            <h2> Desfășurare </h2>
            <h3> Structură </h3>
            <p> Resursele disponibile mai jos completează și nu substituie laboratoare. </p>
          <p>
          <nav>
            <h3> Calendar </h3>
          <aside class="menu supplement">
            <section id="week1">
              <h4> Săptămâna 1 </h4>
              <div class="info">
                <ul>
                  <li class="presentation">
                  <li class="presentation">
                  <li class="lab">
                  <li class="tools">
```

6 matching nodes

6 noduri `` (de tip Element) care au ca valoare a atributului **class** șirul de caractere **"presentation"**

Săptămâna 2

- Curs: **Programare Web – Protocolul HTTP, conținut dinamic prin CGI**

HTTP, câmpuri-antet, mesaj, MIME, pr

- Demonstrații: **Exemple de script-uri CGI** (ar

- Laborator: Foi de stiluri în cascadă (CSS)

- Instrumente: **W3C CSS Validator, Web Deve (Firefox), Chrome Developer Tools, HTTP**

Săptămâna 3

- Curs: **Programare Web. Cookie-uri. Sesiun aplicații. Arhitectura aplicațiilor Web**

programare Web, cookie, sesiune, server de aplicații, MVC

- **Server-side W Modern serve**

Tutoriale: **ASP.NET (exemple), Java (exemple), Perl (exemple), PHP (exen**

- Laborator: De la HTTP la script-uri CGI

- Instrumente: **Aptana Studio, Eclipse, Visual Express, Server2Go, XAMPP**

The screenshot shows a web browser window with a document structure tree on the right. The document is an HTML page with a root element `<document>`. Inside, there is a `<html>` element with attributes `lang="ro"` and `xmlns="http://www.w3.org/1999/xhtml"`. The `<body>` element contains a `<article>` element with attributes `id="top"` and `class="content"`. Inside the `<article>`, there is a `<header>` element, followed by a `<section>` element with attribute `class="main"`. Inside the `<section>`, there is a `<aside>` element with attribute `class="menu"`, followed by a `<h2>` element with text "Desfășurare", a `<h3>` element with text "Structură", a `<p>` element with text "Resursele disponibile mai jos completează și nu substituie laboratoare.", another `<p>` element, a `<nav>` element, a `<h3>` element with text "Calendar", another `<aside>` element with attributes `class="menu"` and `supplement=""`, and a `<section>` element with attribute `id="week1"`. Inside the `<section id="week1">`, there is a `<h4>` element with text "Săptămâna 1", followed by a `<div>` element with attribute `class="info"`. Inside the `<div class="info">`, there is a `` element. The `` contains four `` elements. The first two `` elements have attribute `class="presentation"`. The third `` element has attribute `class="lab"`. The fourth `` element has attribute `class="tools"`. The FirePath extension's XPath search bar at the top right shows the expression `//*[@class="tools" or @class="presentation"]` and the results at the bottom show 11 matching nodes.

```
<document>
  <html lang="ro" xml:lang="ro" xmlns="http://www.w3.org/1999/xhtml">
    <head>
    <body>
      <article id="top" class="content">
        <header>
        <section class="main">
          <aside class="menu">
          <h2> Desfășurare </h2>
          <h3> Structură </h3>
          <p> Resursele disponibile mai jos completează și nu substituie laboratoare. </p>
          <p>
          <nav>
          <h3> Calendar </h3>
          <aside class="menu supplement">
          <section id="week1">
            <h4> Săptămâna 1 </h4>
            <div class="info">
              <ul>
                <li class="presentation">
                <li class="presentation">
                <li class="lab">
                <li class="tools">
```

11 matching nodes

furnizați semnificația expresiei XPath:
`//*[@class="tools" or @class="presentation"]`

xsl: xpath – axe

Nodul curent (<i>context node</i>)	self
Nodul părinte	parent
Nodurile copil	child
Nodurile descendente	descendant
Nodurile de tip atribut	attribute
Nodurile spațiilor de nume	namespace

xsl: xpath – axe

Nodurile ascendente ancestor

Nodurile ascendente
ori nodul curent ancestor-or-self

Nodurile descendente
ori nodul curent descendant-or-self

5 (19–23 martie)
4 (12–16 martie)
3 (5–9 martie)
2 (27 februarie–2 martie)
1 (20–24 februarie)

Calendar

Săptămâna 1

Supliment: *HTML, CSS, and c*

- Prezentare generală

detalii organizat

- Curs: Spațiul World Wide Web – aspecte a

terminologie, model client/server, hipertext, U

- Laborator: De la limbajul (X)HTML către HTML,

- Instrumente: XAMPP, Portable Apps, Sublin
HTML Validator

Săptămâna 2

- Curs: Programare Web – Protocolul HTTP
conținut dinamic prin CGI

HTTP, câmpuri-antet, mesaj, MIME, pr

Top Window Highlight XPath: `//div/descendant::a` Eval

```
<document>
  <html lang="ro" xml:lang="ro" xmlns="http://www.w3.org/1999/xhtml">
    <head>
    <body>
      <article id="top" class="content">
        <header>
        <section class="main">
          <aside class="menu">
            <h2>Desfășurare</h2>
            <h3>Structură</h3>
            <p>Resursele disponibile mai jos completează și nu substituie laboratoare. </p>
            <p>
            <nav>
              <h3>Calendar</h3>
              <aside class="menu supplement">
                <section id="week1">
                  <h4>Săptămâna 1</h4>
                  <div class="info">
                    <ul>
                      <li class="presentation">
                        <p>
                          <a title="Prezentare în format PDF" href="presentations/web00Introducere.pdf">
                        </p>
                      </li>
                    </ul>
                  </div>
                </section>
              </aside>
            </nav>
          </p>
        </section>
      </article>
    </body>
  </html>
```

37 matching nodes

`//div/descendant::a`

`//div/descendant::*`

Ce returnează fiecare expresie XPath?

▶ XPath:

`/produse/produs/promo/preceding::*`

▶ XML:

Test

Save

Color

Format

Fix

```
<?xml version="1.0" encoding="utf-8"?>
<produse>
  <produs>
    <nume>Ping Uinux</nume>
    <ofertant>http://www.penguin.info</ofertant>
    <promo>Mascota lunii</promo>
  </produs>
  <produs>
    <!-- un soi de portocale albastre -->
    <nume>Blue Ory</nume>
    <descriere />
  </produs>
  <produs>
    <nume>Ceas „inteligent” cu gust de măr</nume>
  </produs>
  ...
</produse>
```

expresia XPath

`/produse/produs/promo/preceding::*`

oferă rezultatul

`<nume>Ping Uinux</nume>``<ofertant>``http://www.penguin.info``</ofertant>`

folosind axa **preceding** accesăm nodurile precedente
aici, instrumentul online de testare disponibil la

<http://www.xpathtester.com/xpath>

xsl: xpath – sintaxa prescurtată

self	≡	.
parent	≡	..
attribute	≡	@
namespace	≡	:

../nume

//child::*[@nr > 2]

xsl: xpath – funcții

Noduri: `id()`, `position()`, `count()`, `name()`,
`namespace-uri()`, `last()`, ...

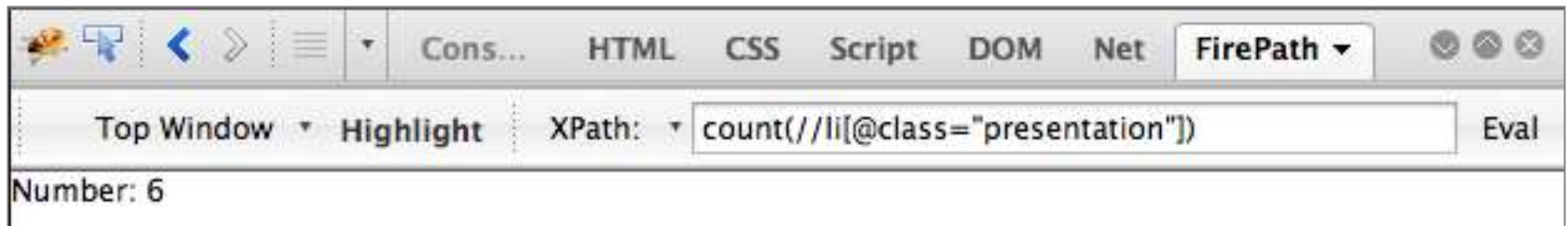
Tipuri de noduri: `node()`, `text()`, `comment()`,
`processing-instruction()`

Șiruri: `concat()`, `starts-with()`, `contains()`, `substring()`,
`string-length()`, `translate()`, ...

Booleeni: `not()`, `true()`, `false()`

Numere: `sum()`, `round()`, `floor()`, `number()`, ...

xsl: xpath – funcții



`count(//li[@class = "presentation"])`

numărarea nodurilor ce satisfac o anumită condiție

xsl: xpath – funcții



`count(//a[@href]) = count(//a)` ► *true*
toate elementele **<a>** includ atributul **href**



`count(//li[@class]) = count(//li)` ► *false*
există elemente **** care nu au specificat atributul **class**

Tehnologii Web

Desfășurare

[evaluare](#) · [regu](#)

Structură

Resursele disponibile mai jos completează și nu su cursuri și laboratoare.

Filmul săptămânal al desfășurării materiei conform **or**:

5 (19—23 martie)

4 (12—16 martie)

3 (5—9 martie)

2 (27 februarie—2 martie)

1 (20—24 februarie)

Calendar

Săptămâna 1

Supliment: *HTML, CSS, and c*

- **Prezentare generală**

[detalii organizat](#)

- Curs: **Spațiul World Wide Web – aspecte a**

terminologie, model client/server, hinertext, III

Cons... HTML CSS Script DOM Net FirePath

Top Window Highlight XPath: `//nav/*/li[position() mod 2 = 1]` Eval

```
<document>
  <html lang="ro" xml:lang="ro" xmlns="http://www.w3.org/1999/xhtml">
    <head>
    <body>
      <article id="top" class="content">
        <header>
          <section class="main">
            <aside class="menu">
              <h2> Desfășurare </h2>
              <h3> Structură </h3>
              <p> Resursele disponibile mai jos completează și nu substituie laboratoare. </p>
            <p>
            <nav>
              <ul class="list">
                <li>
                <li>
                <li>
                <li>
                <li>
              </ul>
            </nav>
            <h3> Calendar </h3>
            <aside class="menu supplement">
```

3 matching nodes

furnizați semnificația expresiei XPath:

`//nav/*/li[position() mod 2 = 1]`

5 (19—23 martie)
 4 (12—16 martie)
 3 (5—9 martie)
 2 (27 februarie—2 martie)
 1 (20—24 februarie)

Calendar

Săptămâna 1

Supliment: *HTML, CSS, and c*

- **Prezentare generală**

detalii organizat

- Curs: **Spațiul World Wide Web – aspecte a**

terminologie, model client/server, hipertext, U

- Laborator: De la limbajul (X)HTML către HTML,

- Instrumente: **XAMPP**, **Portable Apps**, **Sublin**
HTML Validator

Săptămâna 2

- Curs: **Programare Web – Protocolul HTTP**
conținut dinamic prin CGI

HTTP, câmpuri-antet, mesaj, MIME, pr

Top Window Highlight XPath: `//ul[1]/li[last()]/*` Eval

```
<document>
  <html lang="ro" xml:lang="ro" xmlns="http://www.w3.org/1999/xhtml">
    <head>
    <body>
      <article id="top" class="content">
        <header>
        <section class="main">
          <aside class="menu">
            <h2> Desfășurare </h2>
            <h3> Structură </h3>
            <p> Resursele disponibile mai jos completează și nu substituie
              laboratoare. </p>
            <p>
            <nav>
              <ul class="list">
                <li>
                <li>
                <li>
                <li>
                <li>
                <a class="url" href="#week1">
              </ul>
            </nav>
          </section>
        </header>
      </article>
    </body>
  </html>
```

7 matching nodes

găsirea sub-nodurilor elementelor `` aflate pe ultima poziție din primele liste neordonate: `//ul[1]/li[last()]/*`



precizați semnificația expresiei XPath:
`//section[@id and not(contains(@id, "week"))]`

- Curs: Dezvoltarea aplicațiilor Web. Inginerie Web, etape, metodologii, șabloane de proiect
- Laborator: Programare Web
- Instrumente: Cloud9 IDE, BitBucket, GitHul

Săptămâna 5

- Curs: Modelarea datelor. Familia XML și spații de nume
- Demonstrații: Exemple de documente XML
- Laborator: Programare Web
- Instrumente: oXygen /> XML Editor

Remarci

- Pentru condiții optime în vizualizarea prezentărilor **program capabil să proceseze documentele**
- Conținutul acestei pagini poate fi oricând alterat, în mod prealabilă, trebuind considerat drept *work in progress*

Ultima actualizare: 14 martie 2012 • versiunea XVI (1997—2012)
 Situl Web al disciplinei "Tehnologii Web" este disponibil sub licență [Creative Commons](#).

The screenshot shows a web browser window with a FirePath search overlay. The search path is `//*[contains(text(), "XML")]`, and it has found 4 matching nodes. The XML document structure is as follows:

```

<section id="week1">
<section id="week2">
<section id="week3">
<aside class="menu supplement">
<section id="week4">
<section id="week5">
  <h4> Săptămâna 5 </h4>
  <div class="info">
    <ul>
      <li class="presentation">
        <p>
          Curs:
          <a title="Prezentare în format PDF" href="presentations/web04ModelareXML-FamiliaXML.pdf">Modelarea datelor. Familia XML și spații de nume</a>
        </p>
        <div class="desc">date semi-structurate, XML, spații de nume, exemple</div>
      </li>
      <li class="demo">
        <p>
          Demonstrații:
          <a title="Arhivă .zip" href="demos/xml/xml.zip">Exemple de documente XML</a>
        </p>
      </li>
    </ul>
  </div>
</section>
</section>
  
```

`//*[contains(text(), "XML")]`

furnizarea tuturor nodurilor care conțin șirul "XML"

5 (19–23 martie)

4 (12–16 martie)

3 (5–9 martie)

2 (27 februarie–2 martie)

1 (20–24 februarie)

Calendar

Săptămâna 1

Supliment: *HTML, CSS, and c*

- **Prezentare generală**

detalii organizat

- Curs: **Spațiul World Wide Web – aspecte a**

terminologie, model client/server, hipertext, UI

- Laborator: De la limbajul (X)HTML către HTML,

- Instrumente: **XAMPP**, **Portable Apps**, **Sublin**
HTML Validator

Săptămâna 2

- Curs: **Programare Web – Protocolul HTTP.**
conținut dinamic prin CGI

HTTP, câmpuri-antet, mesaj, MIME, pr

Top Window Highlight XPath: `//li[@class="presentation"]/*a[@href]` Eval

```
<document>
  <html lang="ro" xml:lang="ro" xmlns="http://www.w3.org/1999/xhtml">
    <head>
    <body>
      <article id="top" class="content">
        <header>
          <section class="main">
            <aside class="menu">
              <h2> Desfășurare</h2>
              <h3> Structură</h3>
              <p> Resursele disponibile mai jos completează și nu substituie laboratoare.</p>
            <p>
            <nav>
              <h3> Calendar</h3>
            <aside class="menu supplement">
              <section id="week1">
                <h4> Săptămâna 1</h4>
                <div class="info">
                  <ul>
                    <li class="presentation">
                      <p>
                        <a title="Prezentare în format PDF" href="presentations/web00Introducere.pdf">
                      </p>
                    </li>
                  </ul>
                </div>
              </section>
            </aside>
          </nav>
        </p>
      </h3>
    </p>
  </h2>
</aside>
</section>
</header>
</article>
</body>
</html>
</document>
```

6 matching nodes

precizați semnificația expresiei XPath:

`//li[@class="presentation"]/*a[@href]`

5 (19–23 martie)

4 (12–16 martie)

3 (5–9 martie)

2 (27 februarie–2 martie)

1 (20–24 februarie)

Calendar

Săptămâna 1

Supliment: *HTML, CSS, and c*

- **Prezentare generală**

detalii organizat

- Curs: **Spațiul World Wide Web – aspecte a**

terminologie, model client/server, hipertext, U

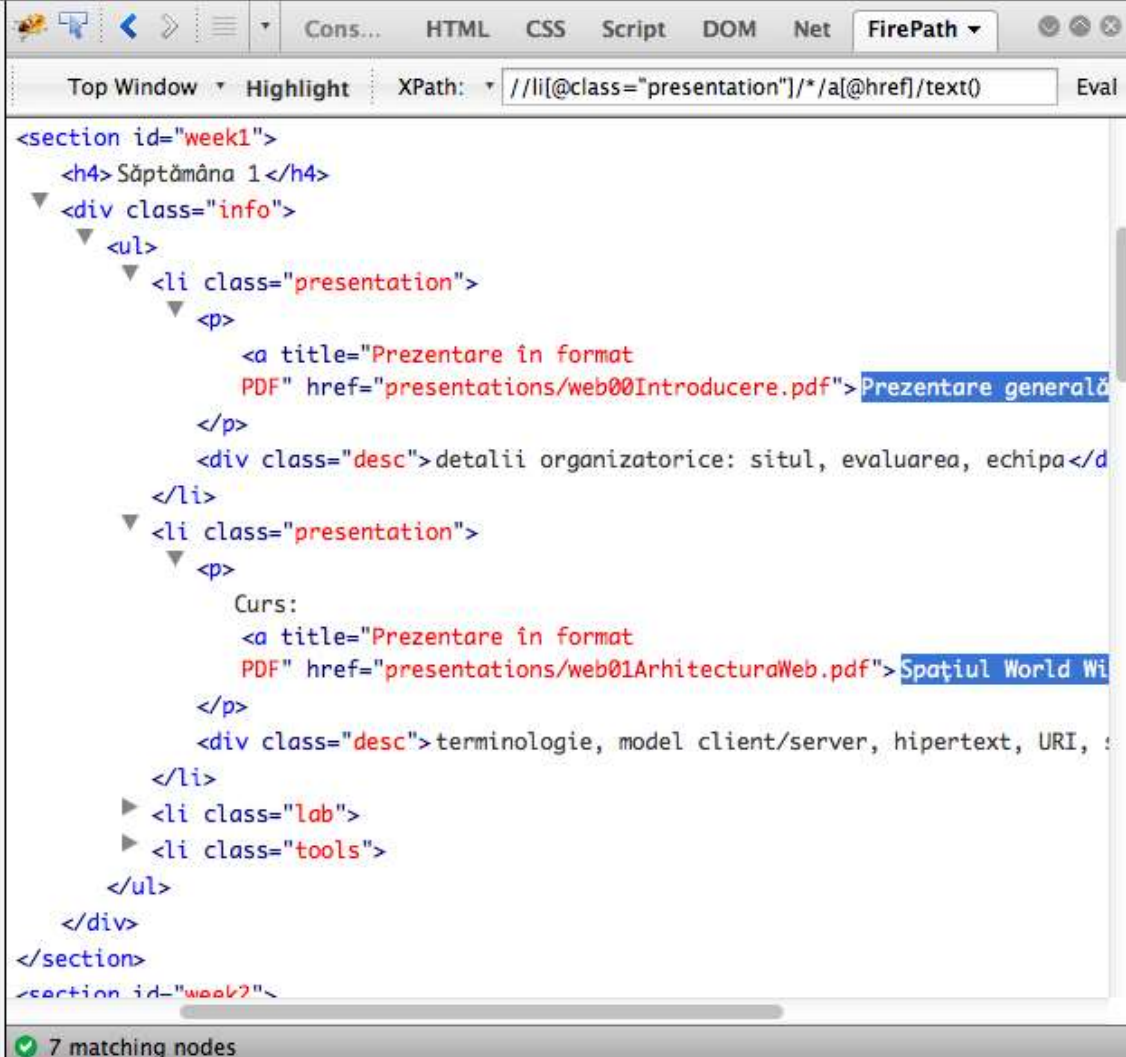
- Laborator: De la limbajul (X)HTML către HTML,

- Instrumente: **XAMPP**, **Portable Apps**, **Sublin**
HTML Validator

Săptămâna 2

- Curs: **Programare Web – Protocolul HTTP**,
conținut dinamic prin CGI

HTTP, câmpuri-antet, mesaj, MIME, pr



```
<section id="week1">
  <h4> Săptămâna 1 </h4>
  <div class="info">
    <ul>
      <li class="presentation">
        <p>
          <a title="Prezentare în format
            PDF" href="presentations/web00Introducere.pdf">Prezentare generală
          </p>
          <div class="desc">detalii organizatorice: situl, evaluarea, echipa</div>
        </li>
      <li class="presentation">
        <p>
          Curs:
          <a title="Prezentare în format
            PDF" href="presentations/web01ArhitecturaWeb.pdf">Spațiul World Wi
          </p>
          <div class="desc">terminologie, model client/server, hipertext, URI,
        </li>
      <li class="lab">
      <li class="tools">
    </ul>
  </div>
</section>
<section id="week2">
```

7 matching nodes

furnizați semnificația expresiei XPath – care-i diferența?

`//li[@class="presentation"]/*[a[@href]/text()]`

xsl: xpath – utilizări pragmatice

Transformarea documentelor XML în alte formate

în conjuncție cu *XSL (Extensible Stylesheet Language)*

XSLT permite generarea de reprezentări
precum HTML, XML sau text obișnuit

XSL-FO facilitează crearea de obiecte de formatare în
vederea tipăririi – *e.g.*, generarea de documente PDF

vezi suplimentul

xsl: xpath – utilizări pragmatice

Procesarea datelor XML

în conjuncție cu *DOM (Document Object Model)*
ori alte metode de prelucrare



detalii în
cursul viitor

xsl: xpath – utilizări pragmatice

Interogarea documentelor XML

în conjuncție cu XQuery

context: servere de baze de date (native XML)

vezi suplimentul

xsl: xpath – programare

libxslt (C, cu portări pentru alte limbaje)

lxml și py-dom-xpath – biblioteci Python

REXML (Ruby)

modul XML cu suport și pentru XPath

SimpleXML – procesări cu XPath în PHP

```
define ('FLUX', 'http://www.infoq.com/feed'); // adresa fluxului de știri RSS
define ('XPATH', '/rss/channel/item'); // expresia XPath utilizată

// funcție ce generează o legătură HTML spre resursă,
// oferind inclusiv descrierea ei
function genereazaLegatura ($url, $titlu, $dataPub = "", $desc = "") {
    return "<section class='news'><p><a href=\"\$url\" title=\"\$titlu\">$titlu</a>
        ($dataPub)</p><div>" . $desc . "</div></section>";
}

try {
    $xml = @simplexml_load_file (FLUX); // încarcăm documentul XML
    // baleiăm însemnările (aici, elementele <item> din RSS)
    foreach ($xml->xpath (XPATH) as $stiri) {
        echo genereazaLegatura ($stiri->link, $stiri->title,
                                $stiri->pubDate, $stiri->description); }
    }
catch (RuntimeException $e) {
    echo $e->getMessage(); // a survenit o excepție
}
```

```

<?xml version="1.0" encoding="UTF-8"?>
<rss version="2.0" xmlns:dc="http://purl.org/dc/elements/1.1/">
  <channel>
    <title>InfoQ</title>
    <link>http://www.infoq.com</link>
    <description>InfoQ feed</description>
    <item>
      <title>Facebook Open Sources React Native, Year
        and Network Class, Fresco and Nuclide</title>
      <link>http://www.infoq.com/news/2015/03/react-native-fresco-nuclide</link>
      <pubDate>Fri, 27 Mar 2015 12:00:00 GMT</pubDate>
      <description>...</description>
      <!-- alte meta-date de interes,
        exprimate și via vocabularul DCMI (spațiul de nume dc:) -->
      <category>Social Networking</category>
      <category>Programming</category>
      <category>Open Source</category>
      <category>Mobile</category>
      <category>Development</category>
      <dc:creator>Abel Avram</dc:creator>
      <dc:date>2015-03-27T12:00:00Z</dc:date>
    </item>
    <!-- alte noutăți... (i.e., elemente <item>) -->
  </channel>
</rss>

```

structura generală a unui
flux RSS poate fi studiată la
<http://www.rssboard.org/>

Presentation: To the Moon (Sat, 28 Mar 2015 15:00:00 GMT)



Russ Olsen tells the moon landing story and how it has affected the software development.

By Russ Olsen

Article: The Indian Banking Community Cloud (Sat, 28 Mar 2015 10:49:00 GMT)



In this article, authors discuss the Indian Banking Community Cloud (IBCC) that provides cloud-based services exclusively to Indian banks to address the financial sector's growing demand for secure cloud-based services.

By Lalit Sangavarapu, Shakti Mishra, Abraham Williams, G.R. Gangadharan

Agile is Giving Testers More Influence (Sat, 28 Mar 2015 09:55:00 GMT)



There is an evolution going on in testing. It used to be that testing was about confirming to the specification. Testers were often brought in too late and had too little influence, but that is changing now as Cirilio Wortel explained in his talk on the evolution of software testing.

By Ben Linders

Presentation: Customer Insight, from Data to Information (Fri, 27 Mar 2015 19:46:00 GMT)



Thore Thomassen shares from experience how to combine structured data in a DWH with unstructured data in NoSQL, and using parallel data warehouse appliances to boost the analytical capabilities.

By Thore Thomassen

Presentation: Lessons Learned Running a Remote Diary Study (Fri, 27 Mar 2015 17:34:00 GMT)



Adam Parker tells how they planned and ran a diary study, what they did during the 3 weeks of the study, how they analyzed the

Managing the Expectations from Agile (Fri, 27 Mar 2015 15:08:00 GMT)



InfoQ did an interview with Gil Zilberfeld about managing the expectations that organizations have of agile and how to prevent misconceptions, valuable ideas and practices from agile and what the future will bring for agile.

By Ben Linders

Measuring the Business Value in Agile Projects (Fri, 27 Mar 2015 15:04:00 GMT)



Technique of "value points" to determine the value delivered by any software project.

By Savita Pahuja

Facebook Open Sources React Native, Year and Network Class, Fresco and Nuclide (Fri, 27 Mar 2015 12:00:00 GMT)



Following yesterday's announcement of releasing ComponentKit, a framework for creating native UIs declaratively on iOS, Facebook has released a number of other projects as open source during their annual F8 developer conference: React Native, Year Class, Network Connection

rezultatul
procesării

xsl: xpath – programare

xpath (JavaScript) – modul Node.js
vezi și <https://nodejsmodules.org/tags/xpath>

XML::XPath – modul Perl

XPath API – diverse interfețe de programare
oferite de **Apache Xalan** și **Saxon** pentru Java
vezi și **javax.xml.xpath.XPath** din cadrul JAXP

hxt-xpath – pachet Haskell
(*Haskell XML Toolbox XPath library*)

8. WeGame

Sa se conceapa un joc Web multi-player (e.g., de tip RPG) care sa aiba drept scop ilustrarea folosirii in practica si invatarea de catre jucatori a celor mai importante concepte predate la disciplina "Tehnologii Web": protocolul HTTP, cookie-uri, sesiuni Web, mecanismul de functionare a unui server de aplicatii Web, dezvoltarea si invocarea de servicii Web pe baza REST, transferul asincron de date prin Ajax, realizarea de mash-up-uri, aspecte esentiale privind securitatea Web si altele. Pe baza API-ului REST creat, jocul va putea fi jucat si in grup via o platforma sociala precum Facebook, inclusiv pe telefoane mobile.

Resurse suplimentare:

- <http://www.slideshare.net/busaco/web-game-development>
- <http://www.slideshare.net/busaco/proiectarea-jocurilor-electronice>
- <https://developers.facebook.com/>



The screenshot shows a web browser's developer console. The top bar includes tabs for Elements, Network, Sources, Timeline, Profiles, Resources, Audits, and Console. The Console tab is active, showing a command prompt with the XPath query: `> $x("//p[@class='info' and contains(text(), 'joc')]/text()")`. The result is a text block containing a paragraph about a multi-player web game. A red box highlights the XPath query in the console output.

```
> $x("//p[@class='info' and contains(text(), 'joc')]/text()")
[
  "Sa se conceapa un joc Web multi-player (e.g., de tip RPG) care sa aiba drept scop ilustrarea
  folosirii in practica si invatarea de catre jucatori a celor mai importante concepte predate
  la disciplina "Tehnologii Web": protocolul HTTP, cookie-uri, sesiuni Web, mecanismul de
  functionare a unui server de aplicatii Web, dezvoltarea si invocarea de servicii Web pe baza
  RE
  pr
  gr
  ,
  "Folosind una sau mai multe API-uri referitoare la jocuri electronice, se doreste dezvoltarea
  unui joc web de strategie, incluzand micro-management de resurse multimedia."
]
```

utilizarea obiectului **\$x** oferit de consola *browser*-ului Chrome sau Firefox pentru evaluarea expresiilor XPath

xsl: xpath

Instrumente XPath – exemplificări:

FirePath, Firefinder – extensii Firebug pentru formularea de interogări XPath asupra documentelor Web

XPath Helper, Xpath Finder – extensii pentru Chrome

EXPath – inițiativă de utilizare/creare de extensii pentru XPath: <http://expath.org/>

Cum verificăm corectitudinea datelor XML?

necesități

Dorim ca datele modelate via XML
să poată fi regăsite, reutilizate și partajate
între aplicații

necesități

O cerință importantă este de a cunoaște:

elementele/atributele ce pot fi specificate

+

modul lor de structurare

e.g., ordinea, numărul minim/maxim de apariții,...

necesități

O cerință importantă este de a cunoaște:
tipul conținutului

exemplu: „atributul align poate avea valorile
mutual exclusive left, right, center”

necesități

O cerință importantă este de a cunoaște:
ce poate fi considerat valid și ce reprezintă eroare

exemplificare:
elementul `` are conținut vid
(nu va putea include alte elemente)

soluție

Grupuri de indivizi specifică vocabularul XML
(setul de elemente/atribute permise)
și regulile de marcarea – **modelele structurale**

soluție

Grupuri de indivizi specifică vocabularul XML
(setul de elemente/atribute permise)
și regulile de marcarea – **modelele structurale**

uzual, grupurile reprezintă:
o companie, o industrie, persoane împărtășind
interese comune,
dezvoltatori de instrumente de marcarea,
un consorțiu/organizație non-profit

soluție

Modelul structural se aplică unei **clase** de documente XML, în vederea verificării
– via un analizor (procesor, *parser* XML) –
a corectitudinii **instanțelor** de documente
aparținând acelei clase

soluție

Aspecte de interes:

- numirea elementelor/atributelor
- definirea regulilor de utilizare a acestora
- specificarea structurii și conținutului
- definirea constrângerilor
- oferirea unui set de convenții de numire

soluție

Apare necesitatea specificării
unui set de **constrângeri** asociate documentelor
XML, astfel încât datele XML să fie verificate
daca sunt **valide** sau nu din punct de vedere
structural ori al tipului conținutului

soluție

Modalitățile de specificare a constrângerilor
se pot baza pe:

descrieri

reguli

șabloane

soluție

Modalitățile de specificare a constrângerilor
se pot baza pe:
descrieri

„există un element **<utilizator>**
având un atribut **nume** care are conținutul...”


DTD, XML Schema

soluție

Modalitățile de specificare a constrângerilor
se pot baza pe:
reguli

„orice element **<utilizator>** va avea un atribut **nume**,
iar conținutul acestui atribut se va conforma regulii...”

Schematron



vezi
suplimentul

soluție

Modalitățile de specificare a constrângerilor
se pot baza pe:
șabloane – *pattern*-uri

„structura documentelor din clasa ‘utilizator’ trebuie
să se potrivească (*match*) conform următorului șablon...”

RELAX NG

vezi
suplimentul

dtd

Tipuri de documente

Document Type Definition (DTD)

specificare formală a tipurilor de documente
(constituenți + structură)

dtd

Tipuri de documente

Document Type Definition (DTD)

specificare formală a tipurilor de documente
(constituenți + structură)



în spiritul
limbajelor formale

dtd

Documentele XML pot avea sau nu un DTD atașat
intern sau **extern** documentului XML pe care-l validează

dtd

Dacă DTD-ul lipsește, documentul trebuie să respecte un număr minim de constrângeri

- ▶ document **bine formatat** – *well formed*

dtd

Regulile sintactice de specificare
a meta-elementelor DTD provin de la SGML

DTD-ul poate exprima structura conținutului,
indicatorii de apariție, conectorii

detalii: <http://www.w3.org/TR/REC-xml>

dtd

Un DTD specifică un tip de document:

```
<!DOCTYPE element_rădăcină [  
    declarații de elemente, attribute, entități,...  
>
```

dtd

Un DTD specifică un tip de document:

```
<!DOCTYPE element_rădăcină [  
  declarații de elemente, attribute, entități,...  
>
```

```
<!DOCTYPE projects [ ...  
  ...  
>
```

exemplu:
lista proiectelor

dtd: elemente

<!ELEMENT	projects	(project+)>
<!ELEMENT	project	(title, desc?, stud, url?)>
<!ELEMENT	title	(#PCDATA)>
<!ELEMENT	desc	(#PCDATA)>
<!ELEMENT	stud	(#PCDATA)>
<!ELEMENT	url	(#PCDATA)>

meta-
element

nume de
element
(neterminal)

conținut (neterm. & term.)
#PCDATA, EMPTY, ANY
* + ? indicatori de apariție
, | conectori

dtd: elemente

Conținutul unui element poate fi:

șir de caractere (**#PCDATA** – *parsed character data*)

vid (**EMPTY**)

orice alt marcaj/șir de caractere (**ANY**)

dtd: elemente

Un element poate apărea:

o singură dată – numele elementului

de oricâte ori, inclusiv niciodată (*)

macăr o dată (+)

de zero sau unu ori (?)

dtd: elemente

Ordinea apariției elementelor:

secvență (,)

alternativă (|)

dtd: attribute

elementul căruia
îi este asociat

```
<!ATTLIST projects  
update          CDATA          #REQUIRED  
>  
<!ATTLIST project  
class           (A | B | C)     "A"  
>
```

tipul conținutului
CDATA, ID, ENTITY,...
enumerare de valori

mod de apariție
#REQUIRED
#IMPLIED
#FIXED

dtd: attribute

Conținutul unui atribut poate fi:

șir de caractere (**CDATA** – *character data*)

identificator unic (**ID**)

referință la un identificator (**IDREF**)

referință la o entitate (**ENTITY**)

enumerare de valori

dtd: attribute

Un atribut poate fi declarat să apară:

obligatoriu (**#REQUIRED**)

opțional (**#IMPLIED**)

fix – se folosește o valoare *a-priori* specificată
(**#FIXED**)

```

<!ELEMENT html (head, body)>
<!ATTLIST html
    id ID #IMPLIED
    xmlns CDATA #FIXED 'http://www.w3.org/1999/xhtml'>
<!ELEMENT img EMPTY>
<!ATTLIST img
    src CDATA #REQUIRED
    alt CDATA #REQUIRED
    height CDATA #IMPLIED
    width CDATA #IMPLIED>
<!ELEMENT table
    (caption?, (col* | colgroup*), thead?, tfoot?, (tbody+ | tr+))>
<!ELEMENT caption #PCDATA>
<!ELEMENT thead (tr)+>
<!ELEMENT tfoot (tr)+>
<!ELEMENT tbody (tr)+>
<!ELEMENT colgroup (col)*>
<!ELEMENT col EMPTY>
<!ELEMENT tr (th | td)+>

```

fragment din DTD-ul
specificând gramatica
limbajului HTML

dtd: validare

O instanță de document trebuie să aibă atașată o declarație a tipului căruia îi aparține:

```
<!DOCTYPE html PUBLIC  
    "-//W3C//DTD XHTML 1.0 Strict//EN"  
    "http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">  
  
<!DOCTYPE projects SYSTEM "projects.dtd">
```

dtd: validare

Validarea documentului se realizează
prin intermediul unui analizor – procesor, *parser* – XML
(Apache Xerces, libxml, MSXML etc.)
sau utilitar
(OpenSP, <oXygen/>, xmllint, XML Nanny,...)

<!-- DTD definind tipul de document 'cuprins' -->

<!ELEMENT cuprins (coordonator?, capitol+)>

<!ELEMENT coordonator ANY>

<!ELEMENT capitol (nume, autor*, descriere?)>

<!ELEMENT nume (#PCDATA)>

<!ELEMENT autor (#PCDATA)>

<!ELEMENT descriere (#PCDATA)>



<!-- declararea unei entități (cu utilizarea entităților predefinite) -->

<!ENTITY busaco "Sabin Buraga <busaco@infoiasi.ro>";>

<!ATTLIST cuprins

| | | |
|-------|-------|-----------|
| titlu | CDATA | #REQUIRED |
|-------|-------|-----------|

| | | |
|---------|-------|----------|
| editura | CDATA | #IMPLIED |
|---------|-------|----------|

| | | |
|----|-----------------------|---------|
| an | (2004 2005 2007 2009) | "2007"> |
|----|-----------------------|---------|

<!ATTLIST capitol

| | | |
|-------|-------|------------|
| numar | CDATA | #REQUIRED> |
|-------|-------|------------|

<!ATTLIST autor

| | | |
|-------|-------|-----------|
| email | CDATA | #IMPLIED> |
|-------|-------|-----------|

<!DOCTYPE cuprins SYSTEM "cuprins.dtd">

<!-- pentru validare se folosește un DTD extern -->

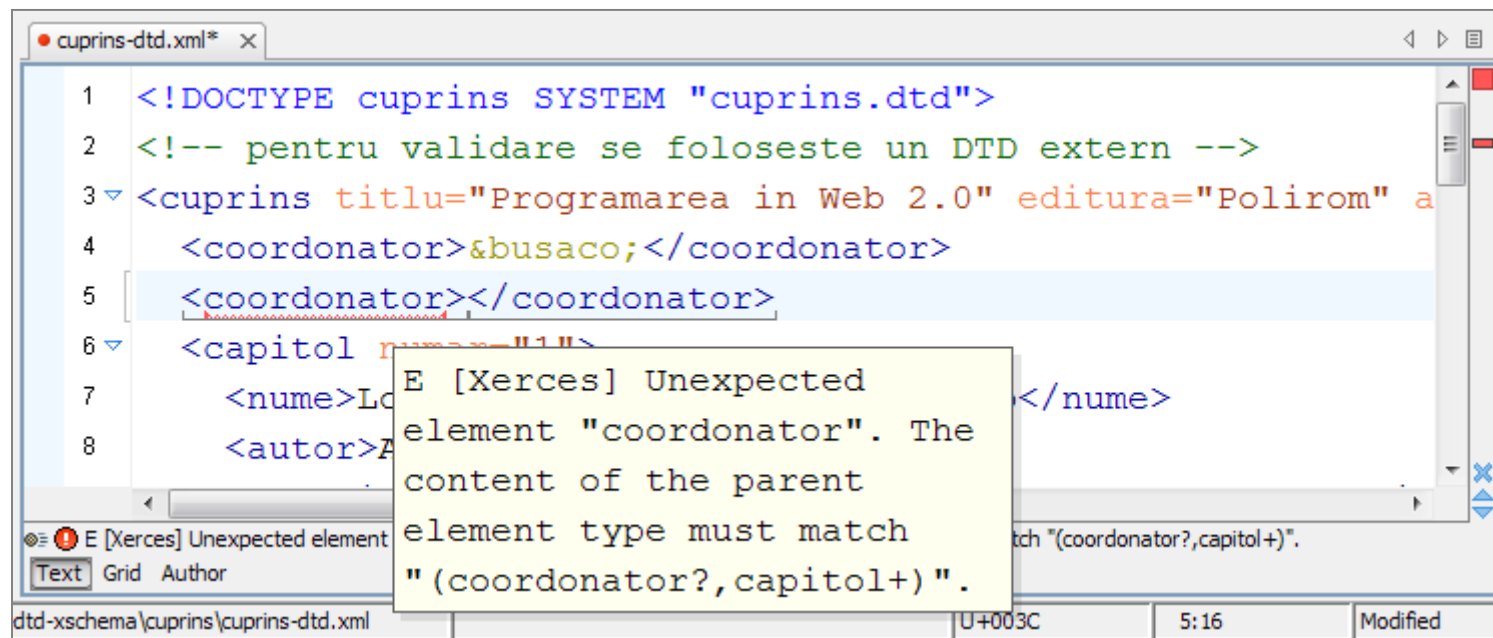
<cuprins titlu="Programarea în Web 2.0" editura="Polirom" an="2007">

<coordonator>&busaco;</coordonator>

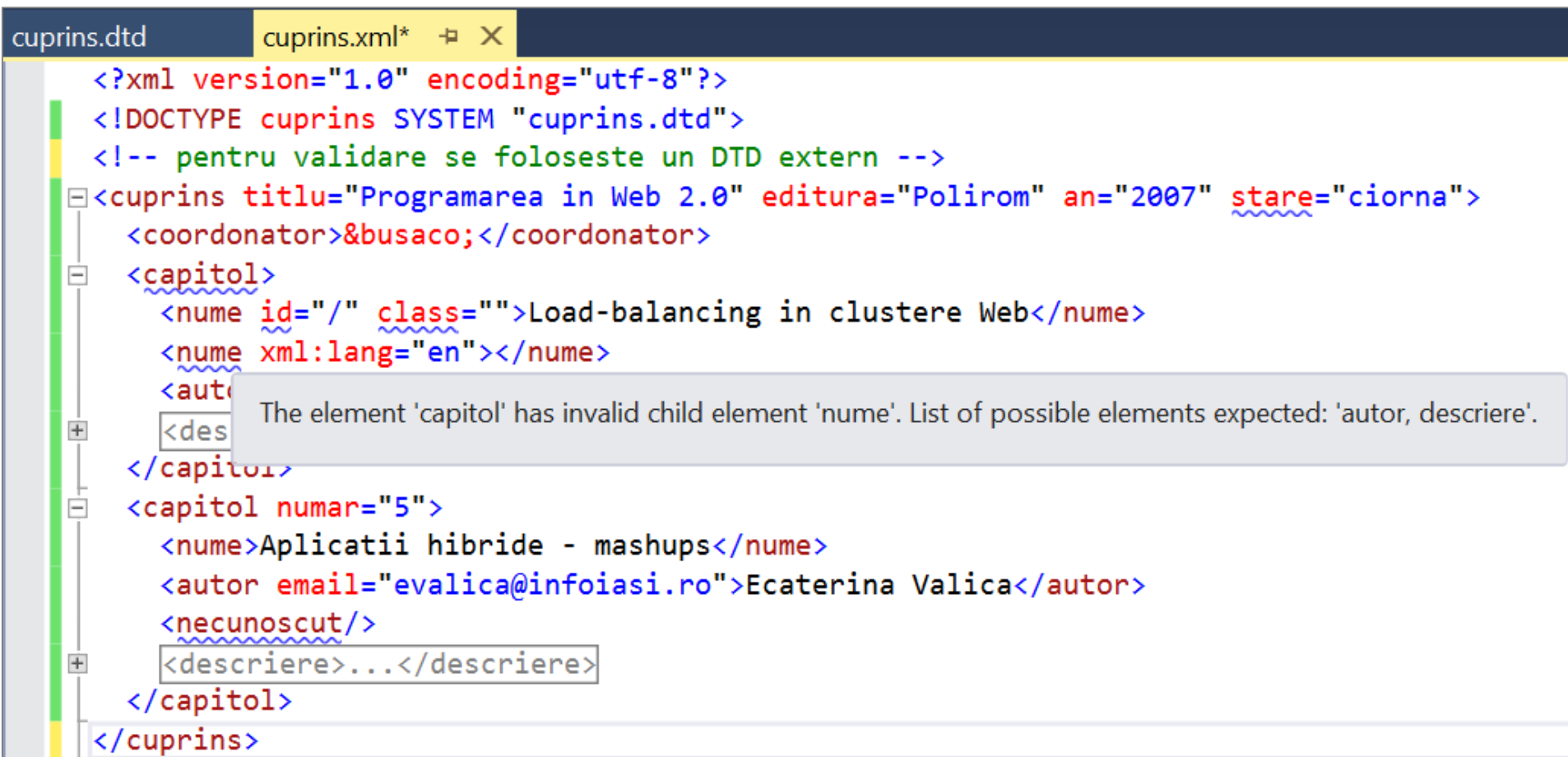
...

</cuprins>

dtd: validare



validarea unui document XML via **Xygen /> XML Editor**
pe baza **Apache Xerces**



```
<?xml version="1.0" encoding="utf-8"?>
<!DOCTYPE cuprins SYSTEM "cuprins.dtd">
<!-- pentru validare se foloseste un DTD extern -->
<cuprins titlu="Programarea in Web 2.0" editura="Polirom" an="2007" stare="ciorna">
  <coordonator>&busaco;</coordonator>
  <capitol>
    <nume id="/" class="">Load-balancing in clustere Web</nume>
    <nume xml:lang="en"></nume>
    <autor>
      <descriere>...</descriere>
    </autor>
  </capitol>
</cuprins>
```

The element 'capitol' has invalid child element 'nume'. List of possible elements expected: 'autor, descriere'.

raportarea erorilor de validare în Visual Studio 2015
recurgând la MSXML

<https://msdn.microsoft.com/en-us/library/ms763742.aspx>

❌ *Line 372, Column 99: value of attribute "align" cannot be "absmiddle"; must be one of "top", "middle", "bottom", "left", "right"*

```
...="RSS" target='_blank'></a>
```

validarea paginilor Web pe baza DTD-urilor
corespunzătoare versiunilor HTML 4.01, XHTML, HTML 5
validator.w3.org

❌ *Line 372, Column 102: required attribute "alt" not specified*

```
...="RSS" target='_blank'></a>
```

The attribute given above is required for an element that you've used, but you have omitted it. For instance, in most HTML and XHTML document types the "type" attribute is required on the "script" element and the "alt" attribute is required for the "img" element.

Typical values for type are type="text/css" for <style> and type="text/javascript" for <script>.

❌ *Line 409, Column 8: required attribute "type" not specified*

```
<script>
```

The attribute given above is required for an element that you've used, but you have omitted it. For instance, in most HTML and XHTML document types the "type" attribute is required on the "script" element and the "alt" attribute is required for the "img" element.

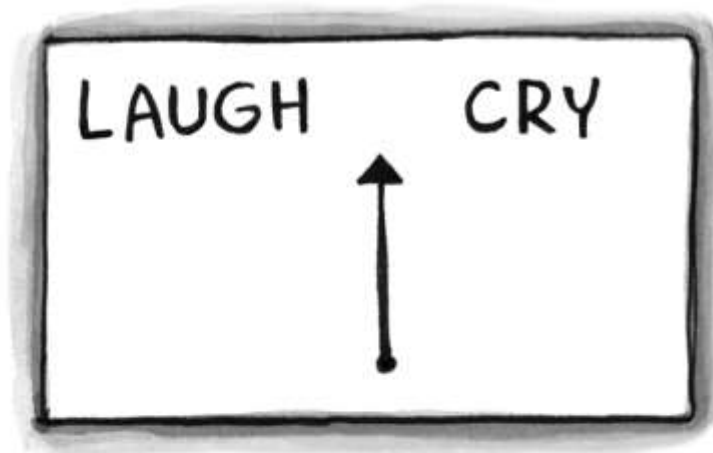
Typical values for type are type="text/css" for <style> and type="text/javascript" for <script>.

rezumat

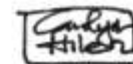
extragerea și validarea datelor XML

//*

limbajul XPath • definirea tipurilor de documente



azilliondollarscomics.com



episodul viitor: **prezentarea proiectului**
în cadrul orelor de laborator (20—24 aprilie 2015)
<http://profs.info.uaic.ro/~busaco/teach/courses/web/web-projects.html>