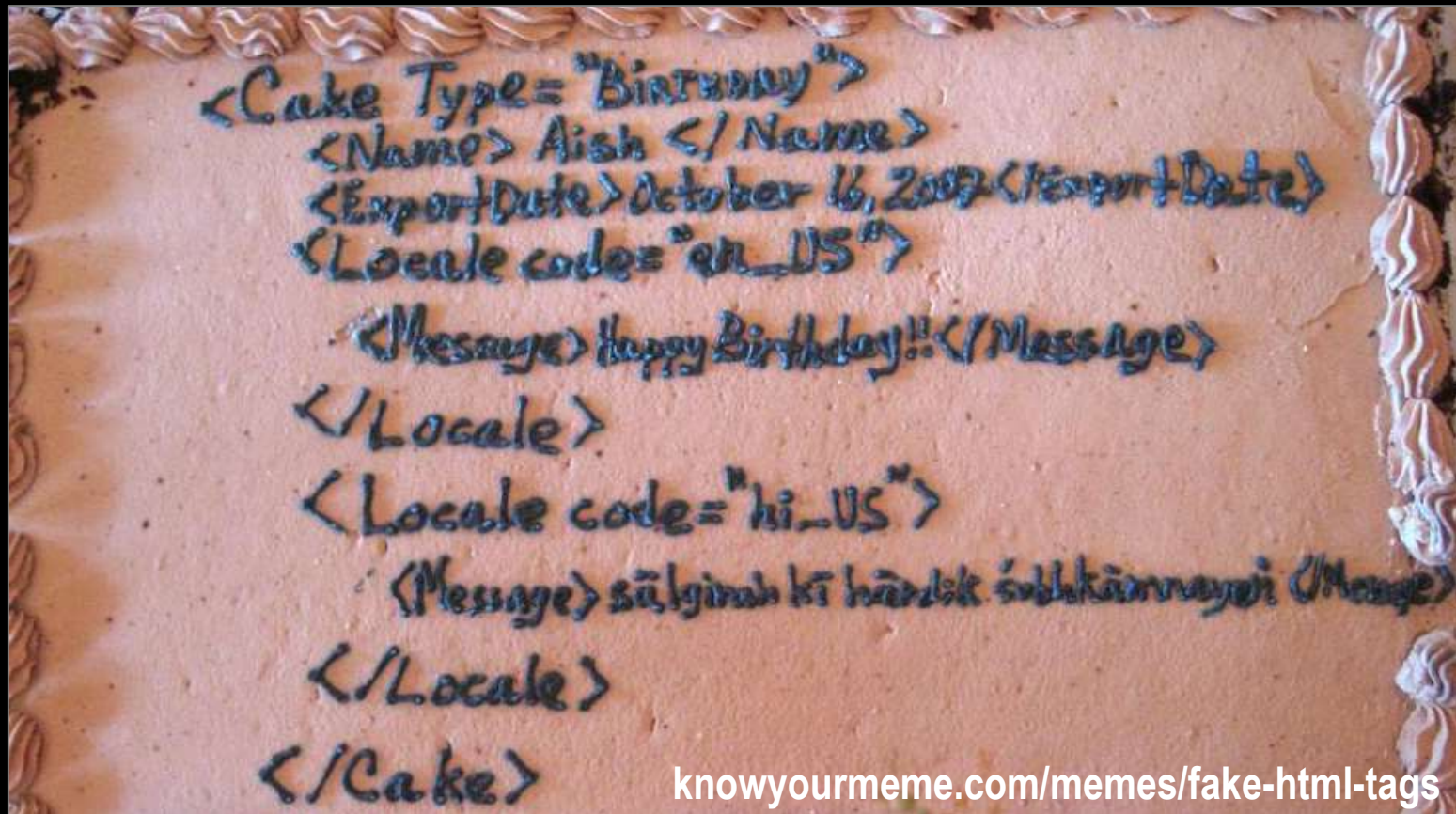


# Tehnologii Web



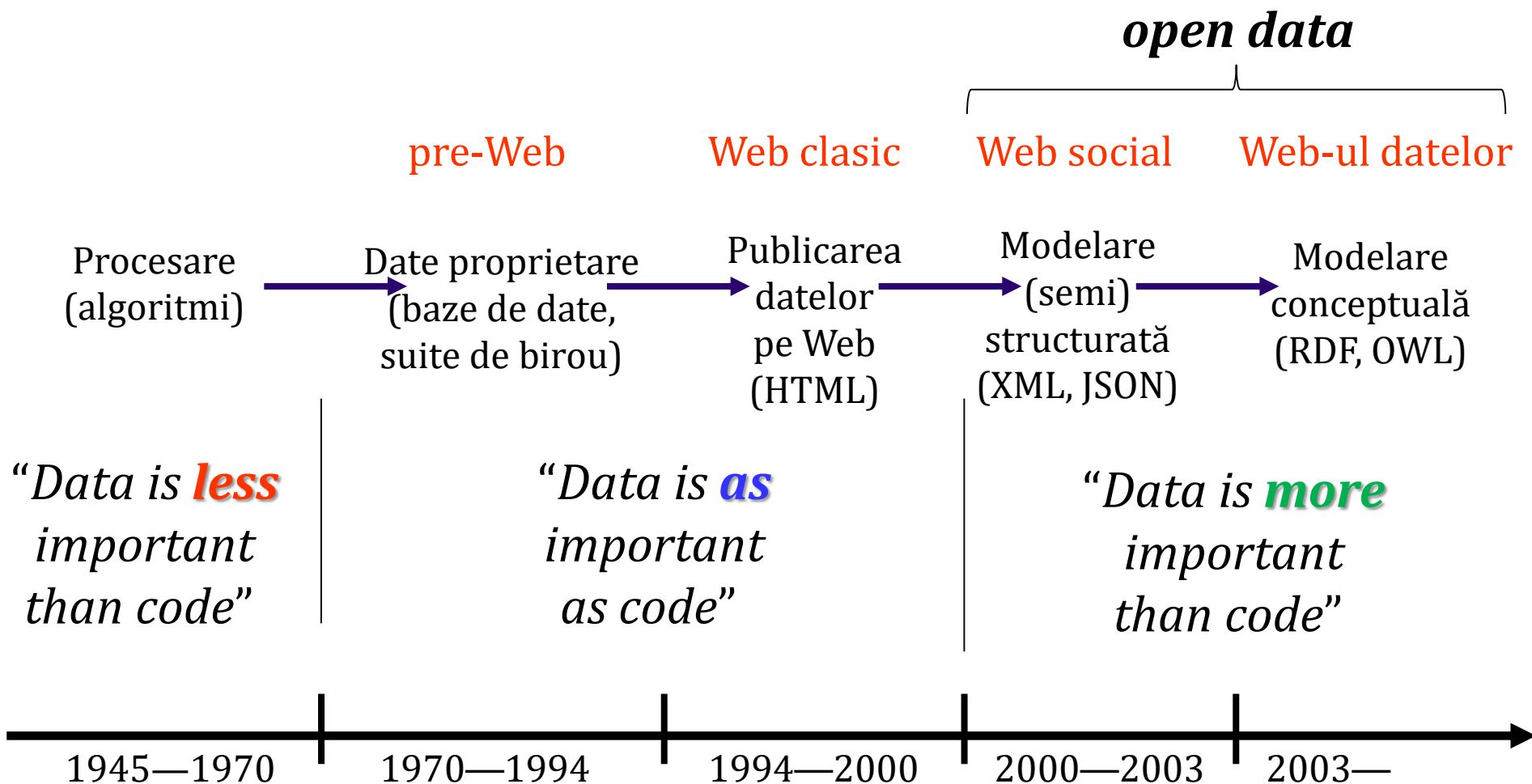
**un model de date pentru Web  
familia XML**

„Esențialul în lumânare nu este ceara  
ce lasă urma, ci lumina.”

**Antoine de Saint-Exupéry**

Cum modelăm (reprezentăm) datele?

# datele: mai importante ca aplicațiile



evoluția conceptului de „dată”  
(adaptare după Daconta *et al.*, 2003)

# preliminarii

Ce model de reprezentare a datelor  
alegem pentru...

stocarea datelor eterogene ce provin din surse multiple?  
informații care evoluează în timp?  
reprezentarea limbajului natural?

# preliminarii

Dorim să modelăm și să procesăm date  
privitoare la

antologii de poezii  
cataloage de produse ale unui *e-shop*  
depozite de rețete gastronomice  
chestionare  
rețele sociale

...

# preliminarii

Necesități:

un limbaj de marcare explicită a informațiilor

datele dorite a fi modelate pot fi practic nelimitate (*unbound*) și necunoscute

nu există *a-priori* un vocabular/schemă comun(ă)

# preliminarii

Necesități:  
datele trebuie să poată fi **auto-explicative**  
(*self-explanatory*)

ce reprezintă triplul ("Sabin", "Buraga", 30374) ?



# preliminarii

Necesități:

modelul adoptat să se preteze arhitecturilor  
navigaționale actuale, bazate pe hipertext

suport pentru specificarea URI-urilor

# preliminarii

Documente:  
formate particulare *vs.* formate generice

# preliminarii

Documente:

formate particulare vs. formate generice

codificare generică ('60):

procedurală – apeluri de proceduri  
bazată pe marcatori (*markups*)

# preliminarii

GenCode – Stanley Rice, Norman Scharpf



GML (*Generalized Markup Language*)

Charles Goldfarb *et al.* (IBM)

definirea formală a tipurilor de documente



SGML (*Standard Generalized Markup Language*)

standard ISO 8879 (1986)

# preliminarii: definiții

Marcaj – adnotare, codare, *markup*

orice acțiune de a interpreta explicit  
o porțiune de text (conținut)

# preliminarii: definiții

Marcaj – adnotare, codare, *markup*

exemple:

semnele de punctuație pentru limbile scrise,  
delimitatorii folosiți în codul-sursă

```

// verificam daca a fost trimisa intr-adevar o imagine
// pe baza tipurilor MIME (Media Types)
$finfo = new finfo(FILEINFO_MIME_TYPE);
if (FALSE === $ext = array_search ($finfo->file($_FILES['img']['tmp_name']),
    array ('jpg' => 'image/jpeg',
           'png' => 'image/png',
           'gif' => 'image/gif'), true)) {
    throw new RuntimeException ('Upload: format incorrect.');
```

marcaje  
speciale

```

// mutam fisierul transferat in directorul cu imagini
// (generam un nume unic pentru fiecare imagine via SHA1)
// detalii la http://php.net/manual/en/function.hash.php
$numeImg = sprintf('%s.%s',
    sha1_file($_FILES['img']['tmp_name']), $ext);
if (!move_uploaded_file ($_FILES['img']['tmp_name'], IMGDIR . $numeImg)) {
    throw new RuntimeException ('Upload: eroare la salvare.');
```

```

// preluam datele EXIF
$exif = exif_read_data(IMGDIR . $numeImg, 0, true);
if (FALSE === $exif) {
    // nu exista date EXIF?
    throw new RuntimeException ('Nu exista date EXIF');
```

# preliminarii: definiții

Limbaaj de specificare (de adnotare, de marcare)  
set de convenții de marcare  
utilizate pentru codificarea datelor



# preliminarii: definiții

Limbaj de specificare (de adnotare, de marcare)  
set de convenții de marcare  
utilizate pentru codificarea datelor

definește mulțimea de marcaje obligatorii,  
modul de identificare și de structurare a marcajelor  
pe baza unei/unor specificații (*i.e.* gramatici)

# xml

## *Extensible Markup Language*

meta-limbaj de marcare derivat din SGML  
standard W3C (1998, 2000, 2004, 2006, 2008)

[www.w3.org/TR/xml/](http://www.w3.org/TR/xml/)

Tim Bray, *XML is 20*, februarie 2018  
[www.xml.com/articles/2018/02/10/xml-20/](http://www.xml.com/articles/2018/02/10/xml-20/)

XML Technologies including XML, XML Namespaces, XML Schema, XSLT, Efficient XML Interchange (EXI), and other related standards.

### XML Essentials

XML is shouldered by a set of essential technologies such as the infoset and namespaces. They address issues when using XML in specific applications contexts.

### Transformation

Very frequently one wants to transform XML content into other formats (including other XML formats). XSLT and XPath are very powerful tools for creating different representations of XML content.

### Processing

A processing model defines what operations should be performed in what order on an XML document.

### Schema

Formal descriptions of vocabularies create flexibility in authoring environments and quality control chains. W3C's XML Schema, SML, and data binding technologies provide the tools for quality control of XML data.

### Query

XQuery (supported by XPath) is a query language for XML to extract data, similar to the role of SQL for databases, or SPARQL for the Semantic Web.

### Internationalization

W3C has worked with the community on the internationalization of XML, for instance for specifying the language of XML content.

### Security

Manipulating data with XML requires sometimes integrity, authentication and privacy. XML signature, encryption, and xkms can help create a secure environment for XML.

### Components

The XML ecosystem is using additional tools to create a richer environment for using and manipulating XML documents. These components include style sheets, xlink xml:id, xinclude, xpointer, xforms, xml fragments, and events.

### Publishing

XML grew out of the technical publication community. Use XSL-FO to publish even large or complex multilingual XML documents to HTML, PDF or other formats; include SVG diagrams and MathML formulas in the output.

# xml: caracterizare

## Marcaje descriptive

```
<para>  
<img />  
<response>  
<Person>  
<tag>
```

# xml: caracterizare

## Tipuri de documente

### *Document Type Definition (DTD)*

specificare formală a tipurilor de documente  
(constituenți + structură)

folosește la verificarea corectitudinii sintactice

# xml: caracterizare

## Independența datelor

suport pe orice platformă hardware/software  
procesoare XML disponibile  
pentru toate limbajele de programare

translatarea/transformarea documentelor

# xml: trăsături

## Meta-limbaj

capabil a defini alte limbaje de marcare  
extinderea marcajelor

portabil  
independent de codificare/limbă via Unicode

# xml: trăsături

Soluție pentru reprezentarea conținutului  
resurselor Web identificate de URI/IRI

asigurarea inter-operabilității (*lingua franca*)

*documents are data*



# xml: constituienți

Prolog (preambul)

Elemente

Atribute

Entități

Secțiuni CDATA

Instrucțiuni de procesare

# xml: prolog

Declarație care specifică versiunea  
și codificarea documentului

```
<?xml version="1.0"..  
encoding="UTF-8"  
?>
```

atribut  
obligatoriu

atribut  
opțional

trebuie să apară o singură dată la începutul documentului

# xml: elemente

Element = componentă structurală (unitate-text)

# xml: elemente

Element = componentă structurală (unitate-text)

nume – identifică un element

sintaxă similară cu cea a identificatorilor de variabile

produs

# xml: elemente

Sintactic, un element este specificat  
via marcatori (*tag-uri*) – de început și de sfârșit

<produs>Ping Uinux</produs>

# xml: elemente

*Case sensitive*

<marcaj> ≠ <Marcaj> ≠ <MARCAJ>

# xml: elemente

Un element poate avea conținut vid

`<produs></produs>`

`<produs />`

# xml: elemente

Un element poate avea conținut vid

<!-- exemplu concret: JSX (folosit de React.js) -->

```
<Form>  
  <Form.Row>  
    <Form.Label />  
    <Form.Input />  
  </Form.Row>  
</Form>
```

[facebook.github.io/jsx/](https://facebook.github.io/jsx/)



# xml: modele ale conținutului

## Model structural

desemnează relațiile dintre elemente:  
secvență, ierarhie, grupare, incluziune

# xml: modele ale conținutului

Elemente imbricate în alte elemente  
(pot conține date textuale și/sau alte elemente)

**<produs>**

Ping Uinix este o mascotă

**<obs>**policromă**</obs>**

care se vinde

**<obs>**foarte rapid**</obs>**.

**</produs>**

# xml: modele ale conținutului

Elemente imbricate în alte elemente  
(pot conține date textuale și/sau alte elemente)

<!-- marcaje HTML5 respectând convențiile XML -->


```
<article>
  <section>
    <ul>
      <li>Oportunitate:
        <strong>Stagii pe bune</strong>
      </li>
    </ul>
  </section>
</article>
```

# xml: modele ale conținutului

Elementele trebuie să fie închise  
și să se împerecheze corect

`<div><q>We don't need no education</div></q>`

greșit!

<?xml version="1.0" ?> ...  preambul

```
<antologie>
  <poem>
    <titlu>...</titlu>
    <strofa>
      <vers>...</vers>
      <vers>...</vers>
    ...
  </strofa>
</poem>
<poem>
  <titlu />
</poem>
<poem>
  <!-- mai multe poeme... -->
</poem>
</antologie>
```

document XML  
modelând o antologie  
de poezii

```
<?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>🍏 cu gust de 🍏</nume>
```

```
  </produs>
```

```
</produse>
```

date semi-structurate



flexibilitate

un posibil catalog de produse în cadrul unui *e-shop*

# xml: attribute

## Atribut

descrie o anumită proprietate (caracteristică)  
a unei apariții particulare a unui element

# xml: attribute

Atributele apar doar în marcajul de început

```
<antologie stare= "ciorna" data="2019-03-26">
```

...

```
</antologie>
```

```
<student matricol="TuPi33">
```

```
  <nume initiala= "I">Tuxy Pinguinescu</nume>
```

```
</student>
```



# xml: attribute

Atributele pot fi specificate în orice ordine

```
<Button text="Ok" id="@+id/un_buton" />
```

≡

```
<Button id="@+id/un_buton" text="Ok" />
```

# xml: attribute

Numele atributelor este *case sensitive*

``

$\neq$

`<img SRC="..." />`

# xml: attribute

Valorile atributelor trebuie să fie delimitate  
obligatoriu de ghilimele sau apostrofuri  
atributele fără valoare nu sunt acceptate

# xml: attribute

```
<form action=procesare.php method="GET">  
  <label for=căutare">Caută:</label>  
<input default type=search placeholder= /></form>
```

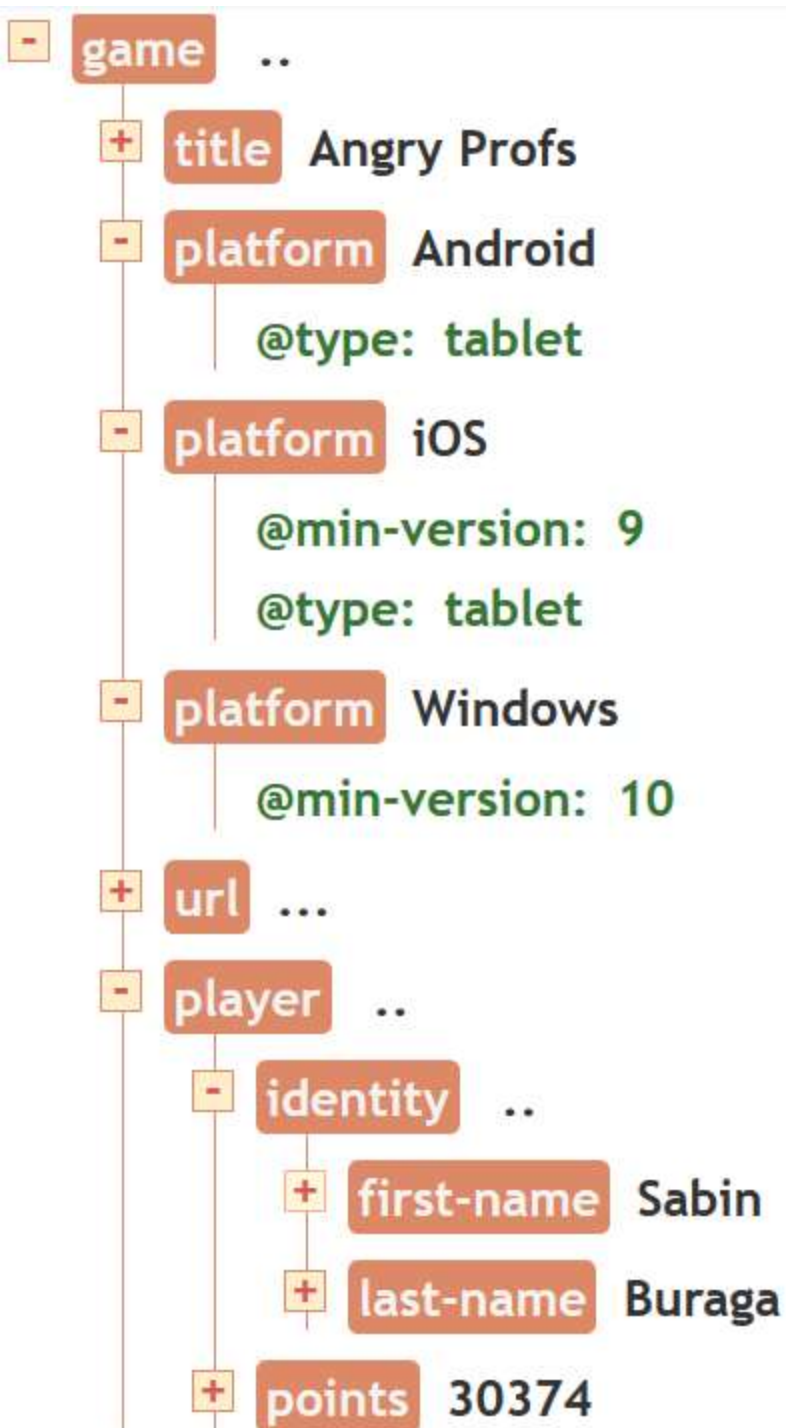
eronat!

```
:</label> Expected attribute's quoted string value.  
type=search placeholder= /></form>
```

```
<game>
  <title>Angry Profs</title>
  <platform type="tablet">Android</platform>
  <platform min-version="9" type="tablet">iOS</platform>
  <platform min-version="10">Windows</platform>
  <url>...</url>
  <player>
    <identity>
      <first-name>Sabin</first-name>
      <last-name>Buraga</last-name>
      <!-- eventual, și alte informații -->
    </identity>
    <points>30374</points>
    ...
  </player>
</game>
```

date XML referitoare la un joc electronic

# vizualizarea structurii ierarhice a datelor XML



# xml: referințe la entități

Scop:

codificare și referențiere a unei părți de document

sintaxă:

**&identificator;**

sau

**&#număr;**

# xml: referințe la entități

Entități predefinite – similare celor de la HTML:

**&lt;** (<) **&gt;** (>) **&amp;** (&) **&quot;** (")

Entități (referințe) de tip caracter:

**&#160;** (*non-breaking space* – **&nbsp;** pentru HTML)

**&#259;** („ă” – setul de caractere ISO-8859-2 și Unicode)

**&#10048;** (simbolul „☼” – Unicode)



# xml: secțiuni

Anumite părți din documente necesită  
procesări speciale

**CDATA** – inhibă procesarea XML

# xml: secțiuni

**<script type="application/javascript">**

```
if (vizite < 10) { // nu e vizitator fidel  
    $("#mesaj").html("<p>Salut!</p>");  
}
```

**</script>**

```
if (visits < 10) { // nu e vizitator fidel  
-----^
```

XML Parsing Error: not well-formed  
Line Number 3, Column 13

## xml: secțiuni

```
<script type="application/javascript">
```

```
/*<![CDATA[*/
```

```
if (vizite < 10) { // nu e vizitator fidel  
    $("#mesaj").html("<p>Salut!</p>");  
}
```

```
/*]]>*/
```

```
</script>
```

procesorul XML nu va interpreta sintaxa codului JavaScript

# xml: instrucțiuni de procesare

Includ informații privitoare la aplicațiile (externe)  
care urmează a fi invocate  
pentru procesarea conținutului

*<?processing-instruction ... ?>*

# xml: instrucțiuni de procesare

```
<script>
```

```
<?php
```

```
    echo "<p>Salut!\n</p>";
```

```
?>
```

```
</script>
```

procesorul XML ar putea invoca, la nivel de server, interpretorul PHP pentru a rula programul

# xml: instrucțiuni de procesare

Exemplificare:

asocierea unei foi de stiluri CSS pentru redarea  
conținutului unui document XML

```
<?xml-stylesheet type="text/css" href="stiluri.css" ?>
```

# xml: procesarea spațiilor

Spațiile albe – *e.g.*, spațiu, TAB, caracterele NL (*New Line*) sau CR (*Carriage Return*) – nu au nicio semnificație

```
<VisualAsset id="obiectObservabil">
```

```
<enabled>true</enabled>
```

```
<zOrder>0</zOrder>
```

```
<Orientation>
```

```
<roll>90</roll>
```

```
<tilt>90</tilt>
```

```
<heading>90</heading>
```

```
</Orientation>
```

```
</VisualAsset>
```

≡

```
<VisualAsset id="obiectObservabil">
```

```
<enabled>true</enabled><zOrder>0
```

```
</zOrder><Orientation><roll>90</roll>
```

```
<tilt>90</tilt><heading>90</heading>
```

```
</Orientation></VisualAsset>
```

marcaje ARML (*Augmented Reality Markup Language*)

# xml: familia

XML (*Extensible Markup Language*)

sintaxa

XML Information Set – Infoset

modelul (abstract) de date XML

XLL (*Extensible Linking Language*)

XLink – legături între documente

XPointer – localizare relativă a resurselor

XSL (*Extensible Stylesheet Language*)

transformări și formatări: XSLT + XSL-FO

XQuery (împreună cu XPath)

interogarea datelor XML



# xml: aplicabilitate

Structurarea/formatarea conținutului  
(formate de prezentare a datelor)

în navigatorul Web: (X)HTML (*Extensible HTML*), HTML5  
[www.w3.org/TR/html/](http://www.w3.org/TR/html/)

formulare electronice: XForms  
[www.w3.org/TR/xforms/](http://www.w3.org/TR/xforms/)

# xml: aplicabilitate

Structurarea/formatarea conținutului  
(formate de prezentare a datelor)

grafică vectorială: SVG (*Scalable Vector Graphics*)

[www.w3.org/Graphics/SVG/](http://www.w3.org/Graphics/SVG/)

lumi tridimensionale: X3D (*Extensible 3D Graphics*)

[www.web3d.org](http://www.web3d.org)

# xml: aplicabilitate

## Reprezentarea diferitelor tipuri de conținut

expresii matematice: MathML

[www.w3.org/Math/](http://www.w3.org/Math/)

date multimedia sincronizate:

SMIL (*Synchronized Multimedia Integration Language*)

[www.w3.org/TR/SMIL/](http://www.w3.org/TR/SMIL/)

# xml: aplicabilitate

## Reprezentarea diferitelor tipuri de conținut

informații vocale: VoiceXML

[www.voicexml.org](http://www.voicexml.org)

informații cartografice: KML (*Keyhole Markup Language*)

[developers.google.com/kml/](http://developers.google.com/kml/)

date hidrologice: WaterML

[www.opengeospatial.org](http://www.opengeospatial.org)

# xml: aplicabilitate

## Reprezentarea diferitelor tipuri de conținut

componente ale interfeței-utilizator:

FXML (JavaFX)

[github.com/mhrimaz/AwesomeJavaFX](https://github.com/mhrimaz/AwesomeJavaFX)

Glade XML pentru GTK+ (GNOME)

[glade.gnome.org](http://glade.gnome.org)

XAML (*Extensible Application Markup Language*)

[docs.microsoft.com/en-us/windows/uwp/xaml-platform/](https://docs.microsoft.com/en-us/windows/uwp/xaml-platform/)

XUL (*Extensible User-interface Language*)

[developer.mozilla.org/docs/Mozilla/Tech/XUL](https://developer.mozilla.org/docs/Mozilla/Tech/XUL)

# xml: aplicabilitate

## Reprezentarea diferitelor tipuri de conținut

documentații: DocBook (*Documentation Book*)  
[docbook.org](http://docbook.org)

informații prelucrate de suite de birou – *e.g.*, Open Office:  
ODF (*Open Document Format*)  
[docs.oasis-open.org/office/v1.2/OpenDocument-v1.2.html](http://docs.oasis-open.org/office/v1.2/OpenDocument-v1.2.html)

# xml: aplicabilitate

## Reprezentarea diferitelor tipuri de conținut

mediatizare (*Web syndication*) – fluxuri de știri:

RSS (*Really Simple Syndication*)

[www.rssboard.org/rss-specification](http://www.rssboard.org/rss-specification)

Atom Syndication Format

[tools.ietf.org/html/rfc4287](http://tools.ietf.org/html/rfc4287)

publicații electronice (*e-books*): EPUB

[idpf.org/epub/](http://idpf.org/epub/)

# xml: aplicabilitate

## Reprezentarea diferitelor tipuri de conținut

informații medicale (EHR – *Electronic Health Records*)

HL7: [www.hl7.org/implement/standards/](http://www.hl7.org/implement/standards/)

afaceri electronice

FpML–*Financial products Markup Language*: [www.fpml.org](http://www.fpml.org)

informații guvernamentale

NIEM–*National Information Exchange Model*: [niem.github.io](http://niem.github.io)



# xml: aplicabilitate



într-un  
curs viitor

## Descrierea serviciilor Web

serializarea datelor transmise conform RPC:

XML-RPC (*XML Remote Procedure Calls*)

descrierea serviciilor:

WSDL (*Web Service Description Language*)

exprimarea protocolului de transfer: SOAP

modelarea proceselor de afaceri via servicii Web:

BPEL4WS

(*Business Process Execution Language for Web Services*)

# xml: aplicabilitate – alte domenii

BeerXML

BDML (*Biological Dynamics Markup Language*)

CAP (*Common Alerting Protocol*)

CML (*Chemical Markup Language*)

COLLADA (*COLLABorative Design Activity*)

DFXML (*Digital Forensics XML*)

GPX (*GPS Exchange Format*)

MEI (*Music Encoding Initiative*)

RTML (*Remote Telescope Markup Language*)

SSML (*Speech Synthesis Markup Language*)

STAR (*Standards for Technology in Automotive Retail*)

TEI (*Text Encoding Initiative*)

# xml: instrumente

## Analizoare (procesoare, *parsers*)

Apache Xerces, Expat, libxml, MSXML,...

## Vizualizatoare și editoare structurale

*browser* Web, Open Office, <oXygen/>, MS Visual Studio

## Formatatoare – *e.g.*, procesoare XSLT/XSL-FO

AH Formatter, FOP, Saxon, Xalan, XEP

## Sisteme de gestiune a bazelor de date XML

BaseX, eXistdb, MarkLogic

Dacă alegem nume de marcatori/attribute  
deja definite de alte limbaje bazate pe XML?

# xml: spații de nume

```
<event uri="https://stagiipebune.ro/">  
  <name xml:lang="ro">Stagii pe Bune</name>  
  <year>2019</year>  
</event>
```

```
<participant>  
  <name uri="mailto:tux@info.uaic.ro">  
    Tuxy Pinguinnesscool</name>  
  <year kind="Bachelor">2</year>  
</participant>
```

# xml: spații de nume

```
<event uri="https://stagiipebune.ro/">  
  <name xml:lang="ro">Stagii pe Bune</name>  
  <year>2019</year>  
</event>
```

conflict!

```
<participant>  
  <name uri="mailto:tux@info.uaic.ro">  
    Tuxy Pinguinnesscool</name>  
  <year kind="Bachelor">2</year>  
</participant>
```

?

# xml: spații de nume

## Spațiu de nume

desemnează un vocabular utilizat pentru calificarea – în mod unic – a elementelor/atributelor XML

# xml: spații de nume

Vocabularul definit – colecție de nume de elemente și attribute, plus maniera lor de structurare – poate fi desemnat de un URI



# xml: spații de nume

Vocabularul definit poate fi desemnat de un URI

atributul **xmlns** specifică acest URI,  
atașând opțional un identificator unic  
fiecărui vocabular folosit

specificație W3C (2009): [www.w3.org/TR/xml-names/](http://www.w3.org/TR/xml-names/)

```
<?xml version="1.0"?>
```

```
<c:calendars xmlns:c="http://www.calendar.info">
```

```
<e:participant xmlns:s="http://www.info.uaic.ro/Students/"  
  xmlns:e="http://www.info.uaic.ro/Events/">
```

```
<s:name>Tuxy Pinguinnescool</s:name>
```

```
<s:year s:kind="Bachelor">2</s:year>
```

```
<c:calendar>
```

```
<e:event xml:id="SpB">
```

```
<e:name xml:lang="ro">Stagii pe Bune</e:name>
```

```
<e:year>2019</e:year>
```

```
</e:event>
```

```
<e:event xml:id="GSoC" />
```

```
</c:calendar>
```

```
</e:participant>
```

```
</c:calendars>
```

fără conflicte!

de studiat exemplele din arhiva  
asociată acestei prezentări

# xml: spații de nume – exemple

Vocabularul XHTML/HTML5: <http://www.w3.org/1999/xhtml>

Vocabularul Atom: <http://www.w3.org/2005/Atom>

Vocabularul modelului conceptual oferit de DBpedia:  
<http://dbpedia.org/ontology/>

Vocabularul documentelor XML *Sitemaps*:  
<http://www.sitemaps.org/schemas/sitemap/0.9>

Vocabularul tipurilor de date XML Schema:  
<http://www.w3.org/2001/XMLSchema>

# xml: spații de nume

Studiu de caz:

includerea în documentele HTML  
a unor construcții provenite din alte limbaje XML

exemplificare:

SVG (*Scalable Vector Graphics*)

[www.w3.org/TR/SVG/](http://www.w3.org/TR/SVG/)

[developer.mozilla.org/docs/Web/SVG](http://developer.mozilla.org/docs/Web/SVG)

[github.com/willianjusten/awesome-svg](https://github.com/willianjusten/awesome-svg)

```
<!DOCTYPE html>
<html xmlns="http://www.w3.org/1999/xhtml">
<head><title>HTML + SVG</title></head>
<body>
<h1>SVG &icirc;ntr-o pagin&#259; Web</h1>

<!-- Elemente și attribute SVG specificate în documentul HTML,
      recunoscute pe baza spațiului de nume SVG -->
<svg width="500" height="200" xmlns="http://www.w3.org/2000/svg">
  <!-- o zonă rectangulară cu colțuri rotunjite -->
  <rect x="50" y="50" rx="7" ry="7" width="450" height="150"
        style="fill: #00CCEE; stroke: #3333CC;"/>
  <!-- conținut textual -->
  <text x="70" y="90" style="stroke: black; fill: gray; font-size: 32pt;">
    SVG direct &icirc;n browser...</text>
  <!-- un cerc galben -->
  <circle cx="400" cy="150" r="33" style="fill: yellow;" />
</svg>

<p>De utilizat un navigator Web oferind suport nativ pentru SVG.</p>
</body>
</html>
```

```

<!DOCTYPE html>
<html xmlns="http://www.w3.org/1999/xhtml">
  <head>
    <title>HTML + SVG</title>
  </head>
  <body>
    <h1>SVG &icirc;ntr-o pagin&#259; Web</h1>

    <!-- Elemente si attribute SVG incluse in documentul HTML,
         recunoscute pe baza spatiului de nume SVG -->
    <svg width="500" height="200"
        xmlns="http://www.w3.org/2000/svg">
      <!-- o zona rectangulara cu colturi rotunjite -->
      <rect x="50" y="50" rx="7" ry="7"
        width="450" height="150"
        style="fill: #00CCEE; stroke: #3333CC;"/>
      <!-- continut textual -->
      <text x="70" y="90"
        style="stroke: black; fill: gray;
              font-size: 32pt;">
        SVG direct &icirc;n browser...
      </text>
      <!-- un cerc galben -->
      <circle cx="400" cy="150" r="33"
        style="fill: yellow;" />
    </svg>

    <p>De utilizat un navigator Web oferind suport nativ
    pentru SVG.</p>
  </body>
</html>

```

## SVG într-o pagină Web

SVG direct în browser...



De utilizat un navigator Web oferind suport nativ pentru SVG.

vizualizarea în navigatorul Web a ilustrației SVG  
via instrumentul **JS Bin**

Web ★

**Safari** pe platforma  
mobilă iOS 12 (iPhone)



redarea conținutului grafic vectorial SVG

**Firefox**  
la nivel de *desktop*  
(Windows 10)

Web ★



```

<svg xmlns="http://www.w3.org/2000/svg"
      xmlns:xlink="http://www.w3.org/1999/xlink" width="8cm" height="3cm">
<title>Grafică vectorială cu SVG</title>
<defs>
  <!-- definim un dégradé (gradient) liniar -->
  <linearGradient id="aGradient">
    <stop offset="33%" stop-color="#ADA" />
    <stop offset="74%" stop-color="#369" />
  </linearGradient>
  <rect id="aRectangle" width="15px" height="15px" rx="2" ry="2" fill="green" />
  <!-- o cale de redare -->
  <path id="aPath" d="M15 50 C10 0 90 0 90 40" />
  <!-- un filtru cromatic -->
  <filter id="visualFilter">
    <feColorMatrix in="SourceGraphic" type="matrix"
      values="0 0 0 0 0
              1 0 1 1 0
              0 1 1 0 0
              0 0 0 1 0" />
  </filter>
</defs>

```

comenzi grafice  
(e.g., M=mută, C=cerc)

aplicarea unui filtru  
[www.w3.org/TR/SVG/filters.html](http://www.w3.org/TR/SVG/filters.html)



```
<!-- o formă rectangulară umplută cu dégradé-ul definit anterior -->  
<rect x="1cm" y="1cm" width="6cm" height="1cm" round="1em"  
      fill="url(#aGradient)" />
```

```
<!-- folosim 4 instanțe ale pătratului definit,  
      plasate la diverse coordonate și având diverse proprietăți grafice -->  
<use x="40" y="40" xlink:href="#aRectangle" />  
<use x="100" y="80" xlink:href="#aRectangle" filter="url(#visualFilter)" />  
<use x="160" y="80" xlink:href="#aRectangle" fill-opacity="0.33" />  
<use x="220" y="80" xlink:href="#aRectangle" />
```

```
<!-- un text redat conform căii specificate -->  
<text fill="red">  
  <!-- de studiat și http://www.w3.org/TR/SVG/fonts.html -->  
  <textPath xlink:href="#aPath">&#10084; Web &#x2605;</textPath>  
</text>  
</svg>
```

# (în loc de) pauză



# xml-infoset

Scop:

specificarea unui model de date abstract  
asigurând inter-operabilitatea tehnologiilor,  
interfețelor de programare (API-uri) și  
aplicațiilor bazate pe XML

recomandare W3C, 2004

[www.w3.org/TR/xml-infoset](http://www.w3.org/TR/xml-infoset)

# xml-info:et: 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 (*depth-first search*)

# 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 aferent 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 dată de **[normalized value]**

# xml-infoset: concepte

Caractere (*characters*)

corespund datelor textuale  
ale conținuturilor elementelor XML

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

proprietatea **[children]** conține datele (text propriu-zis)

# xml-infosec: 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]**

# Element

```
▼ <article>
  ▼ <section itemscope="itemscope" itemtype="http://schema.org/SiteNavigationElement">
    <h2>course</h2>
    ▶ <aside class="special menu"></aside>
    ▼ <nav>
      ▼ <ul>
        ▶ <li class="opt"></li>
        ▶ <li class="opt"></li>
        ▶ <li class="opt"></li>
      </ul>
    </nav>
  </section>
  ▼ <section itemscope="itemscope" itemtype="http://schema.org/SiteNavigationElement">
    <h2>resources</h2>
    ▶ <nav>
  </section>
</article>
```

# Attribute

# Characters

# xml-infoset: concepte

Instrucțiuni de procesare  
(*processing instruction information item*)

Comentarii  
(*comment information item*)

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

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](http://www.w3.org/TR/xpath)

versiuni ulterioare:

XPath 2.0 (2010) – [www.w3.org/TR/xpath20/](http://www.w3.org/TR/xpath20/)

XPath 3 (2014, 2017) – [www.w3.org/TR/xpath-3/](http://www.w3.org/TR/xpath-3/)



# 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 oferit 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  
/game/platform

# xsl: xpath – operatori

Traversare recursivă //

//platform

//div/a

# xsl: xpath – operatori

*Wildcard \**

selectează toate elementele copil ale nodului aflat în contextul curent (*context node*)

**/game/\*/identity  
player/\***

```
<!DOCTYPE html>
<html xmlns="http://www.w3.org/1999/xhtml" lang="ro" xml:lang=
"ro">
  ▶ <head>...</head>
  ▼ <body itemscope itemtype="http://schema.org/WebPage">
    ▶ <header>...</header>
    ▼ <article>
      ▶ <p>...</p>
      ▶ <aside class="menu" lang="en" xml:lang="en">...</aside>
      ▶ <nav>...</nav>
    ... ▶ <div id="past">...</div> == $0
      ▶ <div id="present">...</div>
      ▶ <div id="future">...</div>
      ▶ <section>...</section>
    </article>
    ▶ <footer>...</footer>
  </body>
</html>
```

**`/html/body/*/div`**

evaluarea expresiilor XPath  
în consola *browser*-ului Web  
via obiectul **`$x`**

html body article **div#past**



Console



top



Filter

> \$x ("/html/body/\*/div")

< ▼ (3) [div#past, div#present, div#future] ⓘ

# xsl: xpath – operatori

Nodul curent .

**./player**



# xsl: xpath – operatori

Atribut @

platform/@min-version  
table/@\*

# xsl: xpath – operatori

Spațiu de nume :

svg:\*

```

<!DOCTYPE html>
<html xmlns="http://www.w3.org/1999/xhtml" xmlns:svg="http://www.w3.org/2000/svg">
<head><title>HTML + SVG</title></head>
<body>
<h1>HTML + SVG</h1>

<!-- Elemente si attribute SVG incluse in documentul HTML,
recunoscute pe baza spatiului de nume SVG -->

<svg:svg width="500" height="200">
  <svg:rect x="50" y="50" rx="7" ry="7" width="450" height="150"
  style="fill: #00CC EE; stroke: #3333CC;"/>
  <svg:text x="70" y="90" style="stroke: black; fill: gray; font-size: 32pt;">SVG</svg:text>
  <svg:circle cx="400" cy="150" r="33" style="fill: yellow;" />
</svg:svg>

<p>De utilizat un navigator Web oferind suport nativ pentru SVG.</p>
</body>
</html>

```

evaluarea expresiei **//svg:\*** conduce la obținerea a 4 noduri:

```

<svg:svg xmlns:svg="http://www.w3.org/2000/svg" ...><svg:rect .../>...</svg:svg>
<svg:rect xmlns:svg="http://www.w3.org/2000/svg" .../>
<svg:text xmlns:svg="http://www.w3.org/2000/svg" ...>SVG</svg:text>
<svg:circle xmlns:svg="http://www.w3.org/2000/svg" .../>

```

# xsl: xpath – operatori

Filtru/index [ ]

/game/platform[2]

//li[@class = "lab"]

forma generală: [*expresie*]

```
▼ <section id="week1">
  ▶ <h2>...</h2>
  ▼ <ul>
    ▶ <li class="lecture" itemscope itemtype="http://schema.org/CreativeWork" itemid="#web-lecture0">...</li>
    ▶ <li class="lecture" itemscope itemtype="http://schema.org/CreativeWork" itemid="#web-lecture1">...</li>
    ▶ <li class="bib">...</li>
    ... ▼ <li class="lab"> == $0
      ▼ <p>
        "Laborator: "
        <a href="https://www.w3.org/html/" title="Detalii despre HTML (HyperText Markup Language)">Limbaajul de
        marcare HTML</a>
        " - de studiat "
        ▶ <span itemscope itemtype="http://schema.org/CreativeWork">...</span>
        " și "
        <a href="https://www.html-5-tutorial.com/" title="Un tutorial privitor la HTML">HTML5 - A Tutorial for
        Beginners</a>
        </p>
        ▶ <ul>...</ul>
      
```

html body article div#past section#week1 ul **li.lab**

Console

Filter Default levels ☒ Group similar

```
> $x ('//li[@class="lab"]')
< ▼ (13) [li.Lab, li.Lab, li.Lab, li.Lab, li.Lab, li.Lab, li.Lab, li.Lab, li.Lab, li.Lab, li.Lab, li.Lab, li.Lab]
```

//li[@class = "lab"]

13 noduri **<li>** (de tip Element) având ca valoare a atributului **class** șirul de caractere **"lab"**

# xsl: xpath – operatori

Pentru valori logice și numerice pot fi folosiți  
operatorii uzuali:

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

comparațiile de șiruri sunt *case sensitive*

# 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



```

▼ <section id="week1">
  ► <h2>...</h2>
  ▼ <ul>
    ► <li class="lecture" itemscope itemtype="http://schema.org/CreativeWork" itemid="#web-lecture0">...</li>
    ▼ <li class="lecture" itemscope itemtype="http://schema.org/CreativeWork" itemid="#web-lecture1">
      ...
      ► <aside class="menu">...</aside> == $0
      ► <p>...</p>
      ► <div class="terms" itemprop="keywords" lang="en" xml:lang="en">...</div>

```

html body article div#past section#week1 ul li.lecture **aside.menu**

Styles Event Listeners DOM

⋮ Console

▶ ⓧ | top

Ce rezultat oferă expresiile XPath?

**//section/descendant::aside**

**//section/descendant::\***

```

> $x ('//section/descendant::aside')
< (16) [aside.menu, aside.menu, aside.menu, aside.menu, aside.menu,
  ► aside.menu, aside.menu, aside.menu, aside.menu, aside.menu,
  aside.menu, aside.menu, aside.menu, aside.menu, aside.menu,
  aside.menu]

```

```

> $x ('//section/descendant::*')
< (676) [h2, ul, li.lecture, p, a, span, div.terms, li.lecture,
  ► aside.menu, a, p, a, span, div.terms, li.bib, p, span, a, span,

```

## ▶ XPath:

/game/player/identity/preceding::\*

## ▶ XML:

Test

Save

Color

Format

Fix

```
<?xml version="1.0" encoding="UTF-8"?>
```

```
<game>
```

```
  <title>Angry Profs</title>
```

```
  <platform type="tablet">Android</platform>
```

```
  <platform min-version="9" type="tablet">iOS</platform>
```

```
  <platform min-version="10">Windows</platform>
```

```
  <url>...</url>
```

```
  <player>
```

```
    <identity>
```

```
      <first-name>Sabin</first-name>
```

```
      <last-name>Buraga</last-name>
```

```
      <!-- ... -->
```

```
    </identity>
```

```
    <points>30374</points>
```

```
  </player>
```

```
</game>
```

rezultatul expresiei XPath  
/game/player/identity/preceding::\*

```
<title>Angry Profs</title>
```

```
<platform type="tablet">Android</platform>
```

```
<platform min-version="9" type="tablet">iOS</platform>
```

```
<platform min-version="10">Windows</platform>
```

```
<url>...</url>
```

folosind axa **preceding** accesăm nodurile precedente  
un instrument Web de testare: [www.xpathtester.com/xpath](http://www.xpathtester.com/xpath)

# xsl: xpath – sintaxa prescurtată

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

**../first-name**

**//child::\*[@min-version > 9]**

# 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()`, ...

numărarea nodurilor ce satisfac o anumită condiție

**count(/game/platform[@min-version >= 9])**

```
1  <?xml version="1.0" encoding="UTF-8"?>
2  <game>
3    <title>Angry Profs</title>
4    <platform type="tablet">Android</platform>
5    <platform min-version="9" type="tablet">iOS</platform>
6    <platform min-version="10">Windows</platform>
7    <url>...</url>
```

PROBLEMS   OUTPUT   DEBUG CONSOLE   TERMINAL   XPath Results ▼

XPath Query: count(/game/platform[@min-version >= 9])

[Result]: 2

extensia **XML Tools** pentru Visual Studio Code

# xsl: xpath – funcții

`count(//a[@href]) = count(//a)` ▶ *true*

toate elementele **<a>** includ atributul **href**

`count(//li[@class]) = count(//li)` ▶ *false*

există elemente **<li>** care nu au specificat atributul **class**

```
> $x ('count(//a[@href]) = count(//a)')
```

```
< true
```

```
> $x ('count(//li[@class]) = count(//li)')
```

```
< false
```

```

▶ <section id="week4">...</section>
▼ <section id="week5">
  ▶ <h2>...</h2>
  ▼ <ul>
    ▶ <li class="lecture" itemscope itemtype="http://schema.org/
      CreativeWork" itemid="#web-lecture5">...</li>
    ▼ <li class="demo">
      ▼ <p> == $0
        "Demo: "
        <a href="demos/php/php.zip" title="Arhivă .zip">Exemple
        programe PHP</a>
        " (arhivă "
        <code>zip</code>
        ", 41.5K)"
        </p>
      </li>
    ▶ <li class="bib">...</li>
    ▶ <li class="tools">...</li>

```

**`//*[contains(text(), "Demo")]`**  
furnizarea tuturor nodurilor  
care conțin șirul **"Demo"**

html body article div#past section#week5 ul li.demo **p**

Styles Event Listeners DOM Breakpoints Properties Accessibility

⋮ Console

▶ 🔍 top ▼ Filter Default levels ▼ ☒ Group similar

> \$x ('//\*[contains(text(), "Demo")]')

< ▼ (14) [p, p, p, **p**, p, p, p, p, p, p, p, p, p, p] ⓘ



```
<projects>
  <project class="M">
    <title>Minunat</title>
    <desc>Un proiect minunat!</desc>
    <stud>3</stud>
    <url>http://minunat.info/</url>
  </project>
  <project class="B">
    <title>Minunat--</title>
    <desc>Un proiect care nu e minunat...</desc>
    <stud>2</stud>
  </project>
</projects>
```

media numărului de studenți alocați per proiect:  
**`sum(/projects/project/stud/text()) div count(/projects/project/stud)`**

---

XPath Query: `sum(/projects/project/stud/text()) div count(/projects/project/stud)`

[Result]: 2.5

---



evaluarea expresiilor XPath  
direct în linia de comandă

```
<projects>
  <project class="M">
    <title>Minunat</title>
    <desc>Un proiect minunat!</desc>
    <stud>3</stud>
    <url>http://minunat.info/</url>
  </project>

  <project class="B">
    <title>Minunat--</title>
    <desc>Un proiect care nu e minunat...</desc>
    <stud>2</stud>
  </project>
</projects>
```

instrumentul **xmllint** (folosește **LibXML**)  
oferit de distribuțiile GNU/Linux + macOS

**xmllint projects-dtd.xml**

**--xpath "substring(/projects/project[1]/desc/text(), 4, 7)"**

# xsl: xpath – programare

Exemplificare:

folosirea XPath pentru procesarea unui flux de știri RSS  
(*news feed*) într-un program PHP



de consultat  
arhiva

```
<?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>Protocols are Important: Martin Thompson at QCon London</title>
      <link>https://www.infoq.com/news/2019/03/network-protocols-analysis</link>
      <pubDate>Tue, 19 Mar 2019 03:45:00 GMT</pubDate>
      <description>...</description>
      <!-- alte meta-date de interes,
            exprimate și via vocabularul DCMI (spațiul de nume dc:) -->
      <category>Microservices</category>
      <category>Architecture & Design</category>
      <category>Distributed Systems</category>
      <dc:creator>Jan Stenberg</dc:creator>
      <dc:date>2019-03-19T03:45:00Z</dc:date>
    </item>
    <!-- alte noutăți... (i.e., elemente <item>) -->
  </channel>
</rss>
```

structura generală a unui  
flux RSS poate fi studiată la  
[www.rssboard.org](http://www.rssboard.org)

```
define ('FEED', 'https://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 genLink ($url, $newsTitle, $pubDate = "", $desc = "") {
    return "<section class='news'><p><a href=\"\$url\" title=\"\$newsTitle\"
        >$newsTitle</a> ($pubDate)</p><div>" . $desc . "</div></section>";
}

try {
    $xml = @simplexml_load_file (FEED);          // încărcăm documentul XML
    // baleiăm însemnările (aici, elementele <item> din RSS)
    foreach ($xml->xpath (XPATH) as $news) {
        echo genLink ($news->link, $news->title,
            $news->pubDate, $news->description);
    }
}
catch (RuntimeException $e) { echo $e->getMessage (); }
```

# InfoQ News

resources of interest regarding software development

**[Presentation: Would You Have Clicked on What We Would Have Recommended?](#)** (Wed, 20 Mar 2019 00:00:00 GMT)



Peter B. Golbus describes recent work on the offline estimation of recommender system A/B tests using counterfactual reasoning techniques.

*By Peter Golbus*

**[Presentation: Planting the Service Design Seeds at CBC](#)** (Tue, 19 Mar 2019 21:00:00 GMT)



Hira Javed discusses the increasing adoption of the service design discipline, and CBC's adventures in embracing this approach.

*By Hira Javed*

**[Xiaoya Wei on Airbnb's Knowledge Graph](#)** (Tue, 19 Mar 2019 09:05:00 GMT)



Airbnb's knowledge graph encodes information about their inventory and the world in a graph structure. This structure is based on a hierarchical taxonomy. InfoQ caught up with Xiaoya Wei, software engineer at Airbnb, regarding their knowledge graph and how it helps them categorize their inventory and deliver travel context to their users.

*By Anthony Alford*

**[Protocols are Important: Martin Thompson at QCon London](#)** (Tue, 19 Mar 2019 03:45:00 GMT)



The protocols we use should be studied and practiced more, they are really important in many aspects, Martin Thompson claimed in his presentation at QCon London 2019, where he first looked back at the evolution of mankind and argued that protocols is the most significant human discovery, and then did a critical analysis of the protocols and ideas we use today.

*By Jan Stenberg*

**[Next.js 8 Static Site Framework Adds Serverless Support](#)** (Mon, 18 Mar 2019)

**[Amazon Alexa Enables an Ambient Computing Lifestyle](#)** (Mon, 18 Mar 2019 15:10:00)

**[TC39 Finalizes the Feature Set for ECMAScript 2019](#)** (Mon, 18 Mar 2019 14:16:00 GMT)

**[ImageSharp: Funding an Open Source Project](#)** (Mon, 18 Mar 2019 11:31:00 GMT)

document HTML generat pe baza datelor RSS

# Cum verificăm corectitudinea datelor XML?

# 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 `<img>` are conținut vid  
(nu va putea include alte elemente)

# 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

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 recurg la:

**descrieri** – DTD, XML Schema

**reguli** – Schematron

**șabloane** – RELAX NG

amănunte în cartea S. Buraga, *Tehnologii XML*:

[www.slideshare.net/busaco/sabin-buraga-tehnologii-xml](http://www.slideshare.net/busaco/sabin-buraga-tehnologii-xml)

# dtd

## Tipuri de documente

### *Document Type Definition*

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 în specificația XML: [www.w3.org/TR/REC-xml](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

nume de  
element  
(neterminal)

## 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

conținut (neterminali / terminali)  
#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 (\*)

măcar 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          (M | B)          "M"
>
```

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* stabilită (**#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 de document 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 utilitar  
(exemple: **OpenSP**, **<oXygen/>**, **xmllint**,...)  
sau analizor – procesor, *parser* – XML  
(**Apache Xerces**, **libxml**, **MSXML** etc.)

apelat din linia de comandă sau invocat prin program

*e.g.*, pentru PHP se folosește metoda **DOMDocument::validate**  
[www.php.net/DOMDocument.validate](http://www.php.net/DOMDocument.validate)

••  
vezi arhiva  
cu exemple

```
<projects>
  <project class="D">
    <title>Super</title>
    <title/>
    <desc>Un super proiect!</desc>
    <stud atribut="suplimentar">
      1
    </stud>
    <url>http://www.super.ro/</url>
  </project>
</projects>
```

```
> xmllint projects-dtd.xml --valid
```

projects-dtd.xml:32: element project: validity error :  
**Value "D" for attribute class of project is not among  
the enumerated set**

```
  <project class="D">
```

^

projects-dtd.xml:36: element stud: validity error :  
**No declaration for attribute atribut of element stud**

```
    <stud atribut="suplimentar">1</stud>
```

^

projects-dtd.xml:38: element project: validity error :  
**Element project content does not follow the DTD,  
expecting (title , desc? , stud , url?),  
got (title title desc stud url )**

```
  </project>
```

^

validarea unui document XML cu utilitarul **xmllint**

**Error** Attribute `border` not allowed on element `span` at this point.

From line 4, column 1829; to line 4, column 2042

```
en"><span height=36 alt="Google Translate" border=0
```

```
style="background:url (https://www.gstatic.com/images/brand...
```

```
/1x/googlelogo_color_92x36dp.png) 0 0 no-repeat; width:110px;height:40px;display:inline-
```

Attributes for element `span`:

[Global attributes](#)

**Error** The `select` element cannot have more than one selected `option` descendant unless the `multiple` attribute is specified.

From line 4, column 12239; to line 4, column 12264

```
on><option SELECTED value=es>Spanish</opti
```

**Error** No space between attributes.

At line 4, column 13142

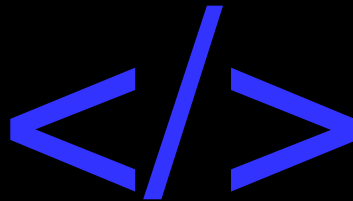
```
nguages"class="jfk-button-stan
```

**Error** Attribute `autocapitalize` not allowed on element `textarea` at this point.

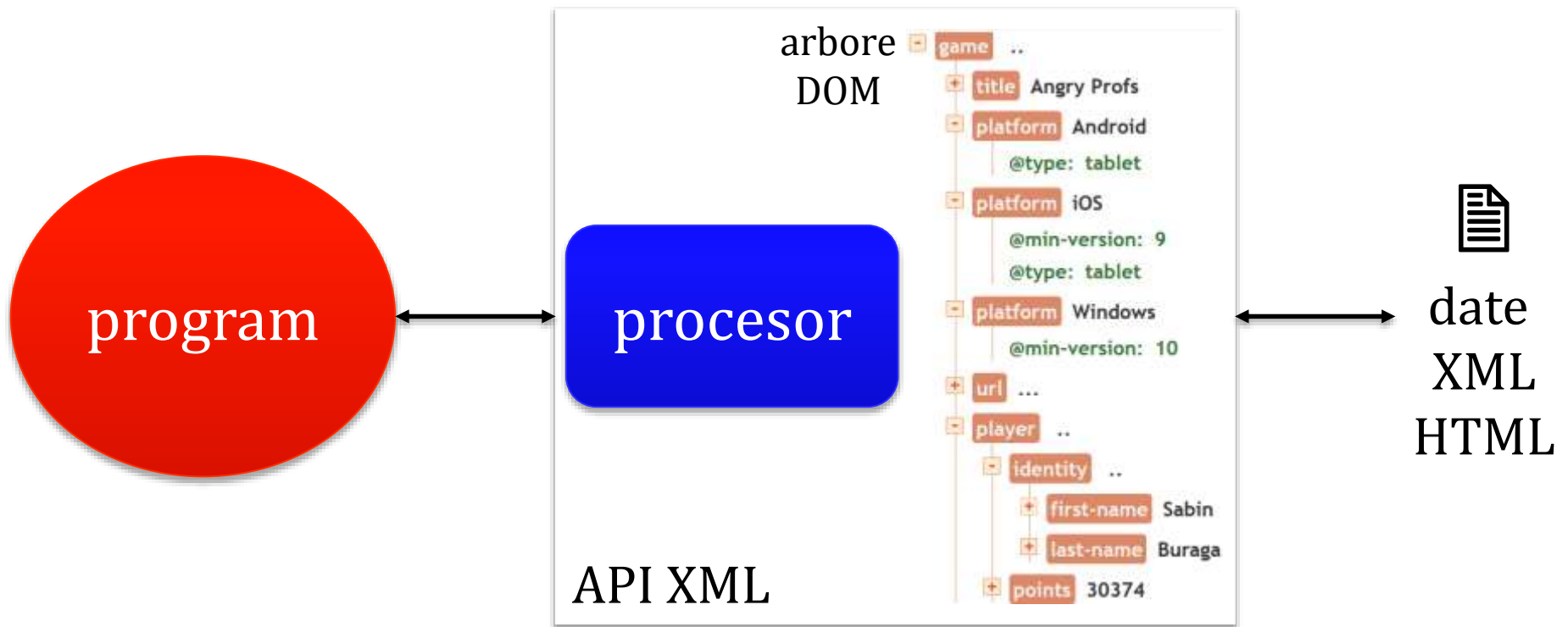
validarea paginilor Web pe baza DTD-urilor  
corespunzătoare versiunilor HTML – [validator.w3.org](https://validator.w3.org)

# rezumat

## modelarea datelor prin XML



caracterizare, utilizări, spații de nume XML,  
acces via XPath, validări DTD



episodul viitor:

**procesarea documentelor XML/HTML cu DOM**