



Rosa Arruabarrena, Jose Ángel Vadillo LSI, UPV/EHU

1

Iturri batzuk

- http://www.w3.org/XML/
- http://www.w3schools.com/xml/default.asp
- http://xml.coverpages.org/xml.html
- http://en.wikipedia.org/wiki/XML
- http://es.wikipedia.org/wiki/XML
- https://www.w3.org/DOM/
- [MS2006] An Introduction to XML and Web Technologies, A. Moller y M. Schwartzbach, Ed. Addison Wesley

HTML noraezean (ikus G3)



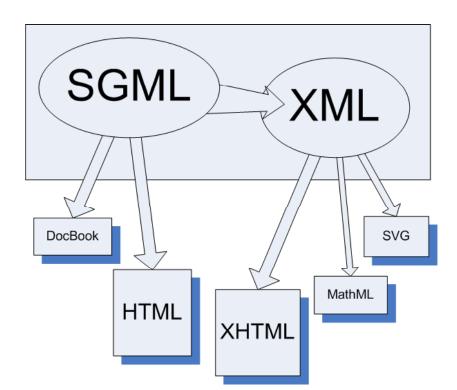
- HTML: SGMLen oinarriturik, bere helburua edukiak deskribatzea da eta internet bidez partekatzea
 - <title>, <address>, <quote>,...
- Formatua emateko eta dokumentuak (edukiak) aurkezteko lengoaian bihurtzen da, web-arakatzailearen bidez erakusten direnak:
 - , , ...
- Baina ez hedagarria

3

W3Cen (World Wide Consortium) erantzuna



- XML: SGMLren azpimultzo bat da edukiak deskribatzeko balio duena
- XMLek ez du ezer egiten. XMLek, informazioa egituratzeko, metatzeko eta trukatzeko balio du
- XML ez dago ez hardwareari ez sistemei, ez plataformei lotua,
- XML internet bidezko informazio trukerako deskribapen lengoaia nagusi bihurtu da, batez ere B2B (business to business) ingurunetan.
- XML erabiliz, lengoaia berriak (dialektoak) defini litezke, metalengoaia da: GML, MathML, RSS, SVG, XHTML,...
- eXtensible Markup Language: XMLko etiketen multzoa ez dago aurredefinitua, "norberak" bere multzoa defini lezake





http://cscie12.dce.harvard.edu/lecture_notes/2007-08/20080130/slide26.html

5

Gainera

- Ulergarriagoa da eta ez hain anbiguoa
- Interneten erabil liteke (testua UNICODE)
- Datuak prozesatzen dituzten programen idazketa errazten du (p.e. bilaketetan)
- Maiz, XML-etaz hitz egiten denean, berez, ez da XML soilez ari baino honen inguruko lengoaia multzo bati eta honen inguruko teknologiei egiten zaie erreferentzia

XML inguruko lengoaia batzuk



Abiapuntua: informazio trukea

(informatikoek: txertatu/eguneratu/kontsultatu)

- Egiaztapena egiteko: XML Schema → .xsd
- Informazioa transformatzeko: XMLT → .xsl
 - Soilik eremu batzuk aurkeratu eta aurkeztu
 - Aurkezpen bat edo beste egiteko: XSLFOO
- Datuen kontsulta egiteko: XPath

7

Adibidea: [MS2006]ren zati bat HTML erabiliz



8

```
<html xmlns="http://www.w3.org/1999/xhtml">
<head>
<title>Some recipes used in the XML tutorial.</title>
<link href="style.css" rel="stylesheet" type="text/css" />
</head>
<body>
>
<h1>Beef Parmesan with Garlic Angel Hair Pasta</h1>
<i>Fri, 28 May 04</i>
<u1>
1.5 pounds of beef cube steak
1 onion, sliced into thin rings
0.25 cup of butter
<01>
```

Adibidea: [MS2006]ren zati bat XML erabiliz

```
<?xml version="1.0" encoding="UTF-8" ?>
<rcp:collection xmlns:rcp="http://www.brics.dk/ixwt/recipes">
<rcp:description>Some recipes used in the XML tutorial. </rcp:description>
<rcp:recipe id="r101">
<rcp:recipe id="r101">
<rcp:title>Beef Parmesan with Garlic Angel Hair Pasta</rcp:title>
<rcp:date>Fri, 28 May 04</rcp:date>
<rcp:ingredient name="beef cube steak" amount="1.5" unit="pound" />
<rcp:ingredient name="onion, sliced into thin rings" amount="1" />
...
<rcp:ingredient name="butter" amount="0.25" unit="cup" />
<rcp:preparation>
<rcp:step>Preheat oven to 350 degrees F (175 degrees C). </rcp:step>
```

Bere DTD-ra

9

ELEMENTUAK



 XML dokumentuak testu lauez osaturik daude (formatu gabeak), markak dituzte (etiketak) garatzaileak definituak

```
<izena> Iruri </izena>
```

- Sintaxia: <etiketa> balioa </etiketa>
- Elementu balio gabeak onartzen dira

```
<etiketa></etiketa>
<etiketa/>
```

Adibidez:

```
<izena></izena>
<izena/>
```

Elementuen arteko guraso-ume erlazioak



• Elementu (guraso) batek, beste elementu bat/batzuk (umeak) eduki ditzake

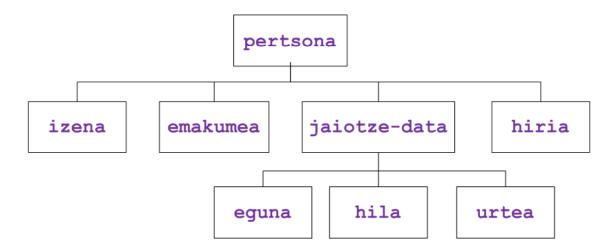
```
<pertsona>
    <izena>Libe</izena>
    <emakumea/>
    <jaiotze-data>
        <eguna>29</eguna>
        <hila>12</hila>
        <urtea>1997</urtea>
        </jaiotze-data>
        <hiria>Donostia</hiria>
</pertsona>
```

11

XML dokumentuko elementu-erroa



 XML dokumentu orok erro bakarra du, eta gainontzeko dokumentuko elementu guztiak haren beheranzkoak dira



• Elementuak dira dokumentuari egitura semantikoa ematen diotenak

Balio mistoa duten elementuak



 Elementu batek eduki mistoa izan dezake, hots, testua eta beste elementuak

 "pertsona" elementuak beste bi elementu ditu, "izena" eta "hiria", eta hauez gain " n jaio da." testua.

13

Oinarrizko ezaugarriak



- XMLeko elementu guztiek itxiera-etiketa izan behar dute
- Elementu guztien izenak case sensitive dira
- XMLeko elementu guztiek egoki azaldu behar dute habiaratuak
- XMLeko dokumentu orok elementu-erro bakarra izan behar du
- Atributuen balioak beti komatxo tartean azaldu behar dute
- XMLn zuriuneak, tabuladoreak, lerro-saltoak, ... ez dira esanguratsuak (Etiketean ezik)
- <!-- iruzkinen sintaxia da hau -->

Oinarrizko sintaxi arauak



- Hizki larriaz nahiz xeheak, zenbakiak, behe marra "_".
 - Ez dira hobesten: puntua "." ez eta erdiko marra "-"
 - Bi puntuen erabilera lagatzen bada era ":" izenen espaziorako (namespace) erreserbatzea aholkatzen da
- Lehengo karakterea letra edo behe marra izan behar du
- Etiketa baten izenaren ondoren zuriune edo lerro saltoa idatz liteke

```
<hiria>Donostia</hiria>
```

 Baina zuriuneak edo lerro saltoak ezin du egon etiketa izenaren aurretik

```
 <
  hiria>Donostia</hiria>
```

15

Adibideak: gaizki idatziak



```
<Hiria>Donostia</hiria>
<hila>12</hil>
<hila>12</hila>
<hiria>Donostia</endhiria>
<_urdina>
<_urdina>
<2kolore>txuria eta urdina</2kolore>
<bi kolore>txuria eta urdina</bi kolore>
< zaletasunak >Zinea, Dantza, Irakurketa</ zaletasunak >
<pertsona><izena>Libe</pertsona></izena>
```

Adibideak: zuzen idatziak



```
<hiria>Donostia</hiria>
<Hiria>Donostia</Hiria>
<hila>12</hila>
<_urdina/>
<kolore2>txuria eta urdina</kolore2>
<bi_kolore>txuria eta urdina</bi_kolore>
<bi-kolore>txuria eta urdina</bi-kolore>
<pertsona>txuria eta urdina</pertsona>
<pertsona>cizena>Libe</pertsona></pertsona></pertsona>
```

Oinarrizko sintaxi arauak



17

- Onartzen badira ere ingelera kanpoko (ñ, ü, à, ê) hizkiak, ez da erabilpena hobesten, beste programekin bateraezintasunak saihestearren.
- Arrazoi berdinagatik, erdi gidoia eta puntua ere ahal bada ez erabili ("," eta "-")

Atributuen sintaxi arauak

- Atributuen izenek elementuen izenen arau berdinak bete behar dituzte
- Elementu baten atributu guztiak desberdinak izan behar dute
- Oker: <datuak x="3" y="4" y="5"/>
- Zuzen: <datuak x="3" y="4" Y="5"/>





19

5 entitate integratuak ditu XMLek

Entity	Entity Reference	Meaning
lt	<	< (less than)
gt	>	> (greater than)
amp	&	& (ampersand)
apos	'	' (apostrophe or single quote)
quot	"	" (double quote)

https://msdn.microsoft.com/es-es/library/ms256190%28v=vs.120%29.aspx

- Me&You idazteko adibidez, erabili Me& You.
- a<b, idazteko a<b. Eta b>c-rentzat aldiz erabili b>c.

🤰 Código fuente de: file:///C:/Users/jiparsar/Desktop/produktuak.xml - Mozilla Fir... 🕒 😐 <u>A</u>rchivo <u>E</u>ditar <u>V</u>er Ay<u>u</u>da <?xml version="1.0" encoding="UTF-8"?> cproduktuak> cproduktua> <izena salneurri="12.56#x20AC">artilezko txanoa</izena> <izena salneurri="12.56#8364">txanoa polarra</izena> file:///C:/Users.../produktuak.xml × **Q** Buscar file:///C:/Users/jiparsar/Desktop/produktuak.xml artilezko txanoa txanoa polarra produktuak.xml ile:///C:/Users/jiparsar/Desktop/produktuak. 🗰 Aplicaciones 🤺 Bookmarks 🌘 Historial » Otros marca This XML file does not appear to have any style information associated w it. The document tree is shown below. ▼oduktuak> <!-- bi produktu daude salgai --> <izena salneurri="12.56#x20AC">artilezko txanoa</izena> <izena salneurri="12.56#8364">txanoa polarra</izena> </produktuak>

XML deklarazioa



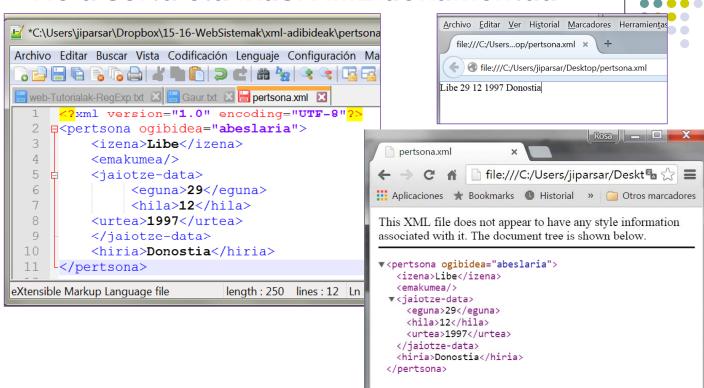
- XMLren deklarazioa ez da prozesaketaren (prozesuaren) agindu bat
- Elementu baten atributu guztiak desberdinak izan behar dute

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

- XML dokumentu batek XML deklarazioa barnean izatea ez da derrigorrezkoa
- Agertzen bada, dokumentuko lehenengo lerroan izan beharko luke eta "<" ikurra lerro horretako lehenengo hizkia
- XML deklarazio sententzia balego, bertsioak agertu behar du.
- Aldiz, encoding eta standalone atributuak hautazkoak dira, eta ezer esan ezean, "UTF-8" eta "yes" dira hurrenez hurren
 - standalone="yes" Fitxategia beste batzuekiko independentea dela adierazten du

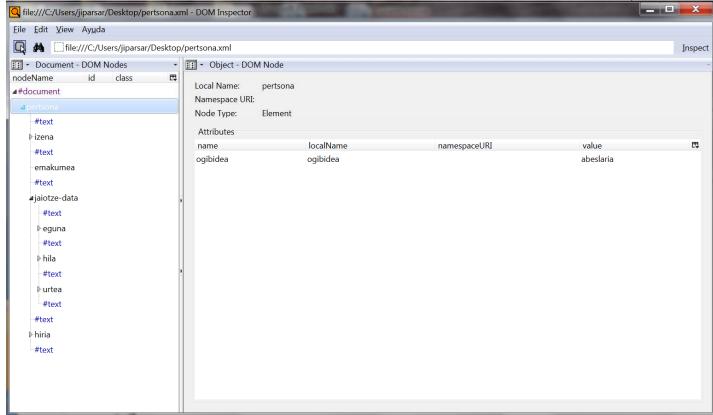
21

Nola sortu eta ikusi XML dokumentua



XML dokumentuaren DOM-a





23

Prozesaketa aginduak



- Prozesaketa aginduak dokumentua prozesatuko duen programari zenbait informazioa emateko da.
- Adibidez: CSS fitxategi bat XML dokumentu bati asoziatu:

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

Adibidez: "estiloa_abereak.css" fitxategiaren edukia

```
izena{color:blue;font-size:40px}
hankak{color:red;font-size:22px}
```

Adibidea: "animaliak.xml"

```
C:\Users\jiparsar\Desktop\animaliak.xml - Notepad++
Archivo Editar Buscar Vista Codificación Lenguaje Configuración Macro Ejecutar TextFX Plugins
Ventana Zen Coding ?
📙 estiloa_animaliak.css 🗵 🔡 animaliak.xml 🗵 📔 pertsona.xml 🗵
      <?xml version="1.0" encoding="UTF-8"?>
      <?xml-stylesheet type="text/css" href="estiloa_animaliak.css"?</pre>
    □<animaliak>
          <animalia>
              <izena>zakur</izena>
              <hankak>4</hankak>
          </animalia>
  8 🛱
              <animalia>
              <izena>ahate</izena>
              <hankak>2</hankak>
          </animalia>
          <animalia>
              <izena>balea</izena>
              <hankak>0</hankak>
          </animalia>
     []</animaliak>
eXtensible Markup Lai length: 350 lines: 16 Ln:1 Col:39 Sel:0 | 0
                                                    Dos\Window UTF-8
```



25

Beraz, XML egitura



- Lehenengo elementua: Bertsioa eta karaktere kodeketa
 - <?xml version="1.0" encoding="UTF-8" ?>
- Erabiltzen diren namespace-n erreferentziak (hautazkoa)
 - <rcp:collection xmlns:rcp="http://www.brics.dk/ixwt/recipes">
- Lengoaiaren deskribapenari erreferentzia (hautazkoa)
 - Dokumentuaren mota eta DTD (Document Type Definition) egiaztapena egiteko
- Elementuak, Atributuak, Iruzkinak, Datuak
 - <rcp:recipe id="r101">
 - Beef Parmesan with Garlic Angel Hair Pasta

DTDren helburua



XML dokumentuaren egitura definitzea da

Zilegi diren elementuen zerrenda eta egitura definitzen ditu

Barruko DTD deklarazioa

27

Egun ere kanpoan definigarria



Kanpo fitxategi batean ere DTD-ak erazagut litezke. Kasu hauetan !DOCTYPEk DTDa duen fitxategiaren erreferentzia jaso beharko du.

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE note SYSTEM "Note.dtd">
<note>
<to>Irune</to>
<from>Lierni</from>
<heading>Jakinarazpena</heading>
<body>Ez ahaztu nitaz asteburuan! </body>
</note>
```





```
<!ELEMENT collection (description,recipe*)>
<!ATTLIST collection xmlns CDATA #FIXED
   "http://www.brics.dk/ixwt/recipes">
<!ELEMENT description (#PCDATA)>
<!ELEMENT recipe
   (title,date,ingredient*,preparation,comment?,nutrition,related*)>
<!ATTLIST recipe id ID #IMPLIED>
<!ELEMENT title (#PCDATA)>
<!ELEMENT date (#PCDATA)>
<!ELEMENT ingredient (ingredient*,preparation)?>
<!ATTLIST ingredient name CDATA #REQUIRED amount CDATA #IMPLIED unit CDATA #IMPLIED>
```

29

[MS2006] errezeten DTDren jarraipena



```
<!ELEMENT preparation (step*)>
<!ELEMENT step (#PCDATA)>
<!ELEMENT comment (#PCDATA)>
<!ELEMENT nutrition EMPTY>
<!ATTLIST nutrition calories CDATA #REQUIRED carbohydrates CDATA #REQUIRED fat CDATA #REQUIRED protein CDATA #REQUIRED alcohol CDATA #IMPLIED>
<!ELEMENT related (#PCDATA)>
<!ATTLIST related ref IDREF #REQUIRED>
```

DTD bidezko errezeten bildumaren egituraren definizioa [MS2006] recipe-rentzat



- Deskribapena. Testua (#PCDATA)
- Errezeten zerrenda (hutsa izan liteke)
- Errezeta
- Izen batez definitua dago, data, osagaien zerrenda, prestatze metodoa, hautazko iruzkina, nutrizio-balioa, erlazionaturik dagoen errezeten zerrenda
- Hautazko atributu bat du, errezetaren identifikadorea dena
- prestatze metodoa
- Pauso-sekuentzia batez osaturik
- nutrizio-balioaren ezaugarriak (atributu bidez emanak) nahi eta ez kaloriak, gantzak, karbohidratoak eta proteinak, eta hautazkoa alkoholari buruzko informazioa

31

DTDa XML dokumentuaren barnean egon liteke (vs external)



```
<?xml version="1.0" encoding="ISO-8859-1" standalone="no" ?>
<!DOCTYPE JUDOKAK [
<!ELEMENT JUDOKAK (JUDOKA)+>
<!ELEMENT JUDOKA (GERRIKOA,IZENA,LORTUA)>
<!ELEMENT GERRIKOA (#PCDATA)>
<!ELEMENT IZENA (#PCDATA)>
<!ELEMENT LORTUA (#PCDATA)>
]>
<JUDOKAK>
<JUDOKAK>
<JUDOKAS
<IZENA>lsabel Fernández Gutiérrez</IZENA>
<LORTUA>1.987</LORTUA>
</JUDOKA>
```

...

Dokumentu zuzen osatuak vs Dokumentu baliodunak



```
<JUDOKAK>
  <JUDOKAS
   <GERRIKOA>beltza</GERRIKOA>
   <IZENA>Isabel Fernández Gutiérrez</IZENA>
   <LORTUA>1.987</LORTUA>
  </JUDOKA>
  </JUDOKAK>
```

Dokumentua zuzen osatu da? (well formed)

Dokumentua balioduna da? (valid)

33

DTDak: helburua? zertarako?



Helburua:

- XML dokumentuaren egitura definitzea da
- Zilegi diren elementuen zerrenda eta egitura definitzen ditu

Zertarako

- Egiaztatzeko jaso diren datuak baliodunak direla
- XML dokumentuek euren formatuen zehaztapena gehiturik eraman dezaten
 - Egun, kanpoko fitxategi batean ere egon liteke DTDa
- Garatzaile o talde desberdinek euren DTD propio komunak defini ditzaten datuen trukea egiteko

Dokumentu zuzen osatuak



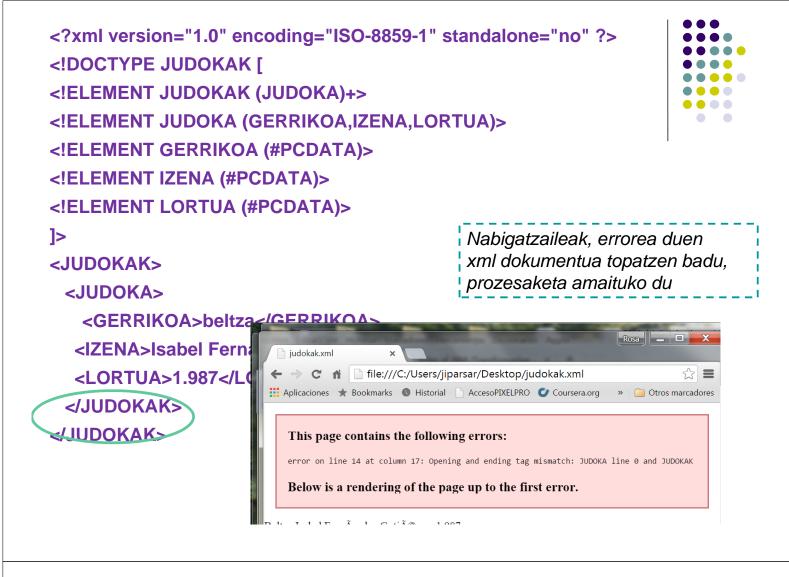
- XMLren zehaztapenak definitzen dituen arauak betetzen dituztenak:
 - Erro bakarra
 - Hasiera-etiketa <... > eta itxiera-etiketa </... > elementu guztientzat.
 - Etiketen habiaraketa zuzena.
 - Case sensitive
 - Atributuen balioak "" artean

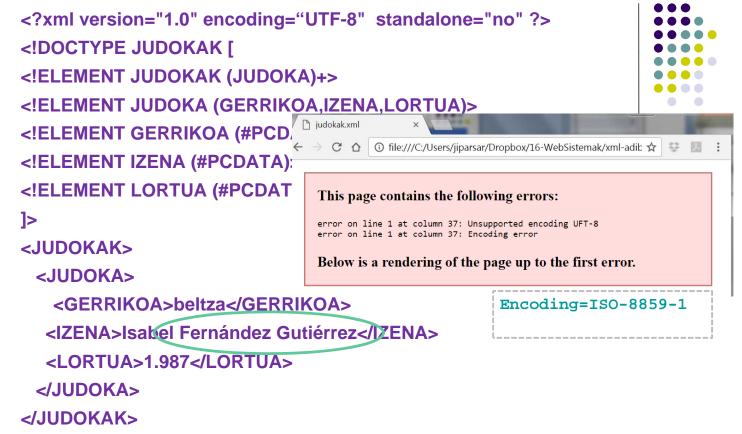
35

Dokumentu baliodunak



- Dokumentu baliodunak derrigorrez zuzen osatua egon behar du
 - Baliodunak dira erlazionaturik dauden DTDra doitzen direnak
 - ...dokumentu zuzen osatuak, ez balioduna izatea gerta liteke





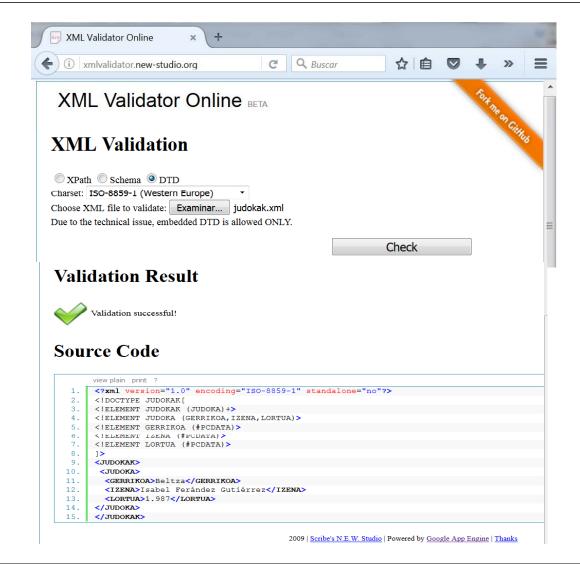
```
<?xml version="1.0" encoding="ISO-8859-1" standalone="no" ?>
<!DOCTYPE JUDOKAK [
<!ELEMENT JUDOKAK (JUDOKA)+>
<!ELEMENT JUDOKA (GERRIKOA,IZENA,LORTUA)>
<!ELEMENT GERRIKOA (#PCDATA)>
<!ELEMENT IZENA (#PCDATA)>
<!ELEMENT LORTUA (#PCDATA)>
1>
<JUDOKAK>
 <JUDOKA>
   <GERRIKOA>beltza</GERRIKOA>
  <IZENA>Isabel Fernández Gutiérrez</IZENA>
  <LORTUA>1.987</LORTUA>
 </JUDOKA>
</JUDOKAK>
```



Online Validated No errors were found

Edo instalatu: "XML developper toolbar" (FireFox).

39





XML dokumentuetako oinarrizko osagaiak

Atzigarriak, kontsultagarri, aldagarriak

1. Elementuak

Oinarrizko blokeak (adibidez: head, body edo table HTMLen

2. Atributuak

- Sartzen den elementuarekiko informazio estra.
- sarrera-etiketan doaz
- Atributuak beti bikotetan deskribatzen dira izena/balioa
- Zenbait ezaugarri azpi-elementuz edo atributu bidez adieraztea diseinu erabakia da
- **3. Entitateak**: < (<), & (&),...
- 4. PCDATA eta CDDATA atalak
 - Eduki mistoa (parsed character data) elementuak+testua, testua (character data)

41

1.Elementuen deklarazioa



- <!ELEMENT elementu-izena category>
 - <!ELEMENT elementu-izena EMPTY>
 - <!ELEMENT br EMPTY>
 - </br>
 - <!ELEMENT elementu-izena ANY>
- <!ELEMENT elementu-izena (elementuaren-edukia)>
 - <!ELEMENT elementu-izena (#PCDATA)>
 - <!ELEMENT elementu-izena (umea-elementu-zena)>





- <!ELEMENT elementu-izena(elementuaren-edukia)>
 - <!ELEMENT elementu-izena (umea-elementu-izena,...)>
 - <!ELEMENT elementu-izena(ume-izena+)>
 - Ume-elementuaren okurrentzia BAT gutxienez
 - <!ELEMENT elementu-izena(ume-izena*)>
 - · Zero edo okurrentzia gehiago ume-elementuena
 - <!ELEMENT elementu-izena(ume-izena?)>
 - Zero edo okurrentzia BAT ume-elementuarena
 - <!ELEMENT elementu-izena(ume-izena|...)>
 - Hautazkoen arteko aukere bat

43

2. Atributuen deklarazioa

- <!ATTLIST elementu-izena atributu-izena atributu-mota default-value>
- Atributu-mota
 - CDATA
 - (en1|en2|...)
 - Zerrendako balio enumeratuen arteko lehenengo biak
 - ID, IDREF, IDREFS, NMTOKEN, NMTOKENS, ENTITY, ENTITIES, NOTATION, ...
- Default-value
 - #REQUIRED, #IMPLIED, #FIXED valor, ...

3. Entitateen deklarazioa



- <!ENTITY entitate-izena "entitate-balioa">
- Adibidea
 - <!ENTITY norkegina"J.A. Vadillo y J.M. Blanco">
 - <!ENTITY irakasgaia-ikasturtea "HAS-12">
 - <refautor>&asignatura-curso;&elaboradospor;</refautor>
- Kanpo entitateen deklarazioa
- <!ENTITY entitate-izena SYSTEM "URI/URL">
 - <!ENTITY writer SYSTEM "http://www.w3schools.com/dtd/entities.dtd">
 - <!ENTITY copyright SYSTEM</pre>
 - "http://www.w3schools.com/dtd/entities.dtd">
 - <author>&writer;©right;</author>

45

DTDen mugak



- Sintaxia ez da XMLerena
- Ez dituzte izenen liburutegiak (namespace) pairatzen
- Ezin dira datu-motak zehaztu, adierazi (adibidez: osokoak, errealak, datak,...)
- Ezin dira adierazi sekuentzia ez ordenatuak
 - ((e1,e2,e3)|(e1,e3,e2)|(e2,e1,e3)|...(e3,e2,e1))
- Ingurunearekiko sentikorrak diren zehaztapenak ezin dira egin
- (Egun DTD-en ordez XML Schema-en erabilpena hobesten da)

XML Schemas



- XML Schema XMLen oinarrituta beste aukera bat da XMLeko DTD-k erabili ordez
 - http://www.w3.org/TR/xmlschema-0/
- XML Schemaren definizio lengoaiek uzten dute :
 - Bokabularioa partekatzea (eskemak, ontologiak,...)
 - XML dokumentuen egitura eta edukiak definitzea
- XML Schema lengoaia XML Schema Definition (XSD) bezala ere ezagutzen da
- EUROPASS: XML Schemas-en aplikazio interesgarria:
 - http://interop.europass.cedefop.europa.eu/data-model/xml-resources/47

XML Schemas (II)



- Dokumentu batean ager litezkeen elementuak
 - "guraso-erlazio" (guraso-ume) dute
 - Agerpen kopuru anitza izan dezakete eta umeen arteko ordena da
- Elementuak hutsik edo testua izan lezakete
- Elementuek eta atributuek datu-motak izan ditzakete
 - Balio lehenetsiak edo aurre-ezarriak

XML Schemas (III)



- Hedagarriak dira
- Adierazkorragoak (DTDak baino)
- XMLen idatzita daude
 - Ez duzu lengoaia berri bat ikasi behar (XMLen aplikazio bat XMLen egina)
 - Eskemak erabiltzeko erremintak (editoreak, egiaztatzaileak,...) XMLkin lan egiteko erabiltzen direnak dira.
 - Eskemak maneia litezke XML-ko DOMa bidez
 - Transfoma litezteke XSLT
 - XML Stylesheets Language for Transformation
- datu-motak pairatzen dituzte
- Izen-espazioak (namespace) pairatzen dituzte
- IKA/GAI "Informazioaren Kudeaketa Aurreratua" irakasgaian sakonkiago jorratzen dira eskemak

XML Schemas, Adibidea



Dokumentuen aurkezpena



- Bi aukera daude
 - 1. CSS-ek HTMLrekin erabiltzen dituen deskribapen berberak erabiltzea
 - XSL-en oinarritzen diren deskribapenak erabiltzea; hots, lengoaia hori erabiliz definitzen estilo orriak erabiliz
 - eXtensible Stylesheet Language
 - http://www.w3.org/Style/XSL/

1.CSS irtenbidea

- Eraginkorra da formatuak eta aurkezpenak deskribatzeko, baina
 - Ez du balio erabakitzeko zeintzuk datu bakarrik erakutsi nahi diren
 - Ondorioz, CSS soilik XML dokumentuekin erabiltzen da dokumentua osorik erakutsi behar denean

51

2.XSL



52

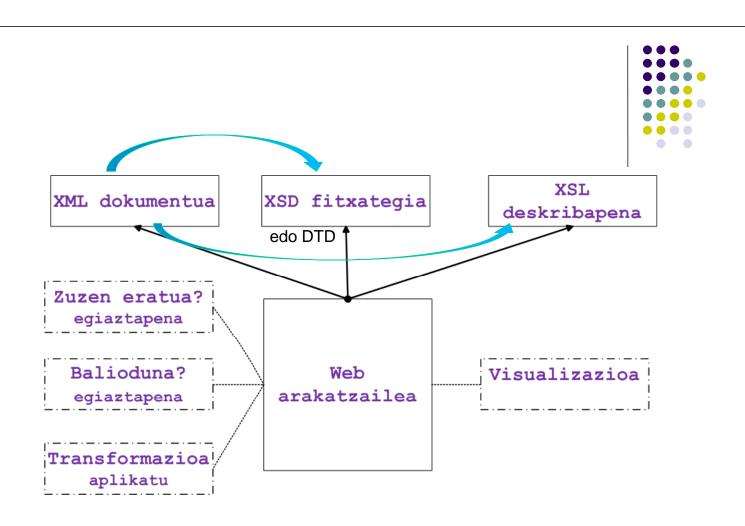
- XSL: eXtensible Stylesheet Language
 - Estilo orriak definitzeko lengoaia hedagarria
- XML sarrera-dokumentu bat transformatzen uzten du irteeradokumentu berri bat sortuz: HTML, PDF, RFT,...
- Osagai desberdinez osatua dago:
 - XSLT: XML dokumentuetan transformazioak zehazteko lengoaiak; hots, XML-ak transformatzeko lengoaia.
 - XL-FO: formateoaren semantika zehazteko XML bokabularioa
 - X-Path: XSLTek erabiltzen duen lengoaia bat da XML dokumentuko atal desberdinak atzitzeko edo erreferentziatzeko

XSL soluzioa (II)



53

- Irtenbide honek ahalbidetzen du
 - Begizten exekuzioa
 - If-then moduko sententzien erabilpenak
 - Konparazio bidezko hautaketak
 - Operazio logikoak
 - Datuen ordenazioa
 - Txantiloien erabilpena



```
<?xml version="1.0" encoding="ISO-8859-1" standalone="no"?>
<?xml-stylesheet type="text/xsl" href="liburuak1.xsl"?>
                                   // transforzio estatikoa
<!DOCTYPE LIBURUAK SYSTEM "liburuak.dtd">
<LIBURUAK>
   <LIBURUA>
      <TITULUA>Speaking in Bones </TITULUA>
      <EGILEA>Kathy Reichs</EGILEA>
                                                    Liburuak.xml
      <PREZIOA>14,99</PREZIOA>
  </LIBURUA>
   <LIBURUA>
      <TITULUA>El domador de leones</TITULUA>
      <EGILEA>Camilla Lackberg</EGILEA>
      <PREZIOA>9,49</PREZIOA>
  </LIBURUA>
<LIBURUA>
      <TITULUA>The Bat</TITULUA>
      <EGILEA>Jo Nesbo</EGILEA>
      <PREZIOA>6,99</PREZIOA>
 </LIBURUA>
                                                                 55
</LIBURUAK>
```

```
<?xml version='1.0'?>
<schema>
  <element name='LIBURUAK'>
  <type>
    <element name='LIBURUA' type='string'/>
      <type>
          <element name='TITULUA' type='string'/>
          <element name='IDAZLEA' type='string'/>
          <element name='PREZIOA' type='fr:currency'/>
      </type>
  </type>
  </element>
                                                      liburuak.xsd
</schema>
                                                       Edo jatorrizkoa
                                                      liburuak.dtd
```

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<xsl:stylesheet version="1.0"</pre>
  xmlns:xsl="http://www.w3.org/1999/XSL/Transform">
<xsl:template match="/">
<HTML>
<BODY>
<xsl:for-each select="/LIBURUAK/LIBURUA">
Titulua: <FONT SIZE="2" COLOR="red" FACE="Verdana">
<xsl:value-of select="TITULUA"/> <BR/>
</FONT>
Egilea: <FONT SIZE="2" COLOR="blue" FACE="Verdana">
<xsl:value-of select="EGILEA"/> <BR/>
Prezioa: <FONT SIZE="2" COLOR="green" FACE="Verdana">
<xsl:value-of select="PREZIOA"/> Euro.<BR/>
</FONT>
                                                Liburuak1.xsl
</xsl:for-each>
</BODY>
</HTML>
</xsl:template>
</xsl:stylesheet>
                                                               57
```

```
irteera
<HTML><BODY>
Titulua:
<FONT SIZE="2" COLOR="red" FACE="Verdana">
Speaking in Bones<BR/>
                                                                                 </FONT>
                                                     🔀 http://localh...sformatu1.php 🗶
Egilea: <FONT SIZE="2" COLOR="blue" FAG
Kathy Reichs<BR/>
                                                      localhost/probak/adibidetransformatu1.php
</FONT>
Prezioa:
                                                  Titulua: Speaking in Bones
<FONT SIZE="2" COLOR="green" FACE="Verd
                                                  Egilea: Kathy Reichs
14,99 Euro. <BR/>
                                                  Prezioa: 14,99 Euro.
                                                  Titulua: El domador de leones
                                                  Egilea: Camilla Lackberg
</FONT>
                                                  Prezioa: 9,49 Euro.
Titulua:
                                                  Titulua: The Bat
<FONT SIZE="2" COLOR="red" FACE="Verda;</pre>
                                                  Egilea: Jo Nesbo
                                                  Prezioa: 6,99 Euro.
Los hombres que no amaban a las mujeres
                                                  Titulua: Los hombres que no amaban a las mujeres
</FONT>
                                                  Egilea: Stieg Larsson
Eqilea:
                                                  Prezioa: 8,54 Euro.
<FONT SIZE="2" COLOR="blue" FACE="Verd;</pre>
Stieg Larsson<BR/>
</FONT>
Prezioa: <FONT SIZE="2" COLOR="green" FACE="Verdana">
8,54 Euro. < BR/>
</FONT>
</BODY></HTML>
```

```
<?xml version="1.0"?>
                                   Beste transformazioa bat
<xsl:stylesheet version="1.0"</pre>
                  xmlns:xsl="http://www.w3.org/1999/XSL/Transform"
<xsl:template match="/">
<HTML><BODY>
<P>NOLA BIHURTU XML DOKUMENTUA HTML TAULA BATEAN</P>
  <TABLE border="1">
    <THEAD><TR><TH>Titulua</TH><TH>Egilea</TH></TR></THEAD>
       <xsl:for-each select="/LIBURUAK/LIBURUA" >
         <TD><FONT SIZE="2" COLOR="red" FACE="Verdana">
         <xsl:value-of select="TITULUA"/> <BR/>
         </FONT>
         </TD>
         <TD>
         <FONT SIZE="2" COLOR="blue" FACE="Verdana">
         <xsl:value-of select="EGILEA"/> <BR/>
         </FONT>
                                            🔀 localhost/xml/adibidetransfor 🗴
         </TD>
                                           ← → C 🔐 🗋 localhost/xml/adibidetransformatu2.php 🖫 🏡 🗏
        </TR>
                                           NOLA BIHURTU XML DOKUMENTUA HTML TAULA BATEAN
    </xsl:for-each>
  </TABLE>
                                                     Titulua
                                                                    Egilea
                                            Speaking in Bones
                                                                  Kathy Reichs
</BODY></HTML></xsl:template>
                                            El domador de leones
                                                                  Camilla Lackberd
</xsl:stylesheet>
                                            os hombres que no amaban a las mujeres Stieg Larsson.
```

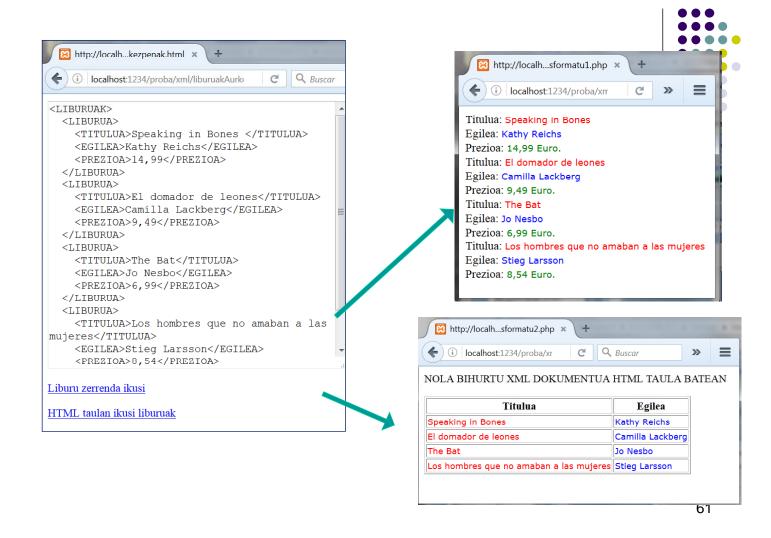
Exekuzio-denboran PHP erabiliz XSL transformazio baten aplikazioa

```
<html><body>
<textarea cols="30" rows="40">
<LIBURUAK>
<LIBURUA>
 <TITULUA>Speaking in Bones
  </TITULUA>
 <EGILEA>Kathy Reichs</EGILEA>
 <PREZIOA>14,99</PREZIOA>
</LIBURUA>
</LIBURUAK>
</textarea>
<a href="adibidetransformatu1.php">
     Liburu zerrenda ikusi</a> 
<a href="adibidetransformatu2.php">
     HTML taulan ikusi liburuak</a>
  </body></html>
```

```
<?php
$xslDoc = new DOMDocument();
$xslDoc->load("liburuakl.xsl");
$xmlDoc = new DOMDocument();
$xmlDoc->load("liburuak.xml");
$proc = new XSLTProcessor();
$proc->importStylesheet($xslDoc);
echo $proc->transformToXML($xmlDoc);
?>
```

```
<?php
$xslDoc = new DOMDocument();
$xslDoc->load("liburuak2.xsl");

$xmlDoc = new DOMDocument();
$xmlDoc->load("liburuak.xml");
$proc = new XSLTProcessor();
$proc->importStylesheet($xslDoc);
echo $proc->transformToXML($xmlDoc);
?>
```



Judoka.xsd



```
<?xml version="1.0" encoding="utf-8"?>
<xs:schema attributeFormDefault="unqualified"</pre>
  elementFormDefault="qualified"
  xmlns:xs="http://www.w3.org/2001/XMLSchema">
  <xs:element name="JUDOKAS">
    <xs:complexType>
      <xs:sequence>
        <xs:element name="JUDOKA">
          <xs:complexType>
            <xs:sequence>
              <xs:element name="CINTURON" type="xs:string" />
              <xs:element name="NOMBRE" type="xs:string" />
              <xs:element name="OBTENIDO" type="xs:decimal" />
            </xs:sequence>
          </xs:complexType>
        </xs:element>
      </xs:sequence>
    </xs:complexType>
  </xs:element>
</xs:schema>
```

XMLeko DOMa



- Document Object Model
 - · Dokumentuko Objektuen eredua
- Web Aplikazioak Programatzeko Interfazea (API) estandarra da (W3C definitua) XML dokumentuak atzitzeko eta maneiatzeko
 - · Lengoaiekiko independentea
- Besteak beste, XML dokumentuak sortzeko balio du, haietan arakatu eta euren edukiak atzitu/aldatu

63

XMLeko DOMeko osagaiak



- Dokumentu bat kargatzen denean nodoz osatutako zuhaitz baten bidez egiten da karga:
 - Nodo bat dokumentuaren datuekin
 - Nodo bat elementu bakoitzeko
 - (testu) elementu bakoitzeko nodo bat balioarekin
 - Nodo bat (atributu,balio) bikote bakoitzeko
 - Nodo bat iruzkin bakoitzeko

Nodo motak (NodeType)



http://www.devguru.com/

NODE_ELEMENT(1)

This node represents an element.

NODE_ATTRIBUTE(2)

This node represents an attribute of an element. Note that it is not considered a child of the element node.

NODE_TEXT(3)

This node represents the text content of a tag.

NODE_CDATA_SECTION(4)

This node represents the CDATA section of the XML source. CDATA sections are used to escape blocks of text that would otherwise be considered as markup.

NODE_ENTITY_REFERENCE(5)

This node represents a reference to an entity in the XML document.

NODE_ENTITY(6)

This node represents an expanded entity.

NODE_PROCESSING_INSTRUCTION(7)

This node represents a processing instruction from the XML document.

NODE_COMMENT(8)

This node represents a comment in the XML document.

NODE_DOCUMENT(9)

This node represents an XML document object.

NODE_DOCUMENT_TYPE(10)

This node represents the document type declaration of the <!DOCTYPE> tag.

NODE_DOCUMENT_FRAGMENT(11)

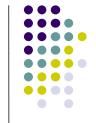
This node represents a document fragment. This associates a node or subtree with a document without actually being part of it.

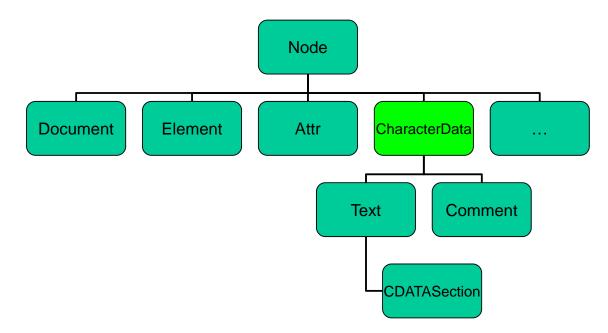
NODE_NOTATION(12)

This node represents a notation in the document type declaration.

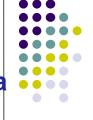
65

XML DOMeko osagaiak





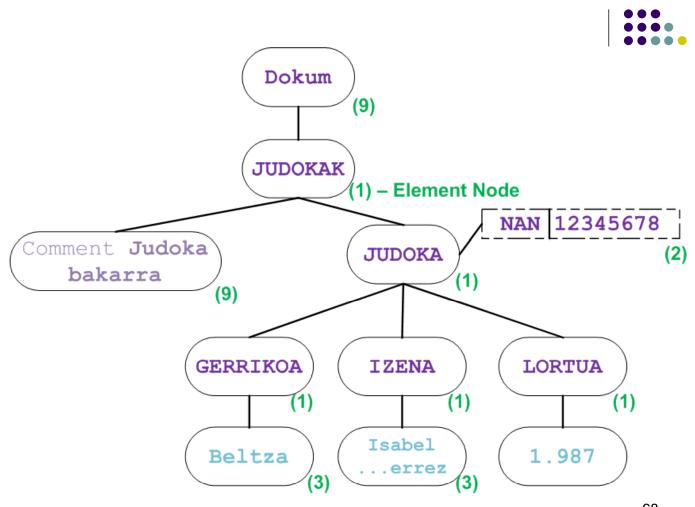
Ariketa: egizu ondorengo dokumentuaren XMLeko DOMa



<?xml version="1.0" encoding="ISO-8859-1" standalone="no"?>

```
<JUDOKAK>
  <!-- Judoka bakarra -->
  <JUDOKA NAN="12345678">
        <GERRIKOA>Beltza</GERRIKOA>
        <IZENA>Isabel Fernández Gutiérrez</IZENA>
        <LORTUA>1.987</LORTUA>
        </JUDOKA>
</JUDOKAK>
```

67





- nodo.nodeName -> nodoaren izena
- nodo.nodeValue -> nodoaren balioa
- nodo.parentNode -> nodoaren gurasoa
- nodo.childNodes -> umeen zerrenda
- nodo.attributes -> atributuen zerrenda

Metodoak

nodo. getElementsByTagName(izena)

emandako izenarekin kointziditzen duten nodoen zerrenda itzultzen du.

nodo.appendChild(nodo-umea)

nodo-ari nodo ume bat txertatzen dio

nodo. removeChild(nodo-umea)
nodo-aren nodo umea ezabatzen du

XML DOMeko beste metodo & propietateak

- doc. documentElement
 Dokumentu bateko erroa itzultzen du
- nodeList.length
 Zerrendak dituen nodo kopurua itzultzen du
- doc. createElement

 (+10 gehiago createNodeType)

 Zehaztutako moteko nodo bat sortzen du
- ...
- http://www.w3schools.com/dom/dom_nodetype.asp

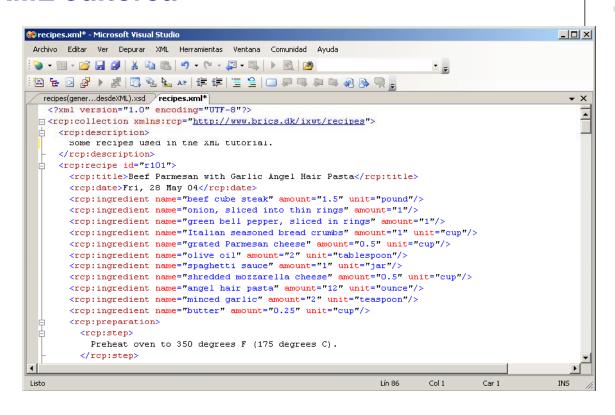


XMLerako erremintak

- XML editoreak
- DTD eta XSD eskema editoreak
- XSL editoreak
- Zuzentzaileak
- Egiaztapena egiteko erremintak:
- DBak XML bihurtzeko erremintak eta alderantziz
- ...

71

XML editorea



XML editorea



- Osaketaren zuzentasuna (well formed)
- Eskemen erlazioa
 - Egiaztapena frogatu
- Transformazioen erlazionatu
 - Froga
- XML dokumentuetatik abiatuz eskemen sorrera

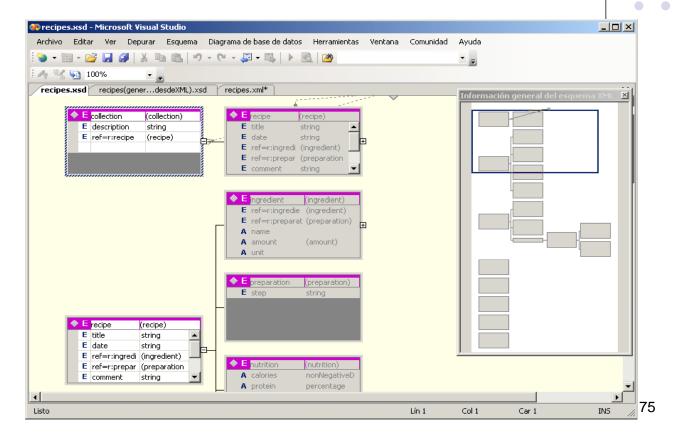
73

DTD editorea

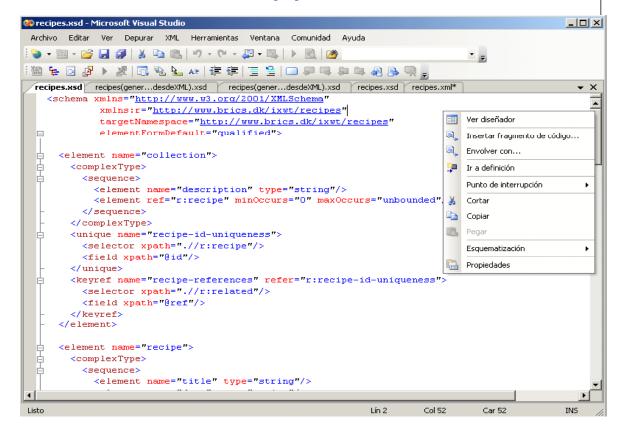


- DTDen egiaztapen sintaktikoa (well formed)
- DTDtik abiatuz haren eskemaren sorrera

Eskemen editorea



Eskemen editorea (II)





• • • • •

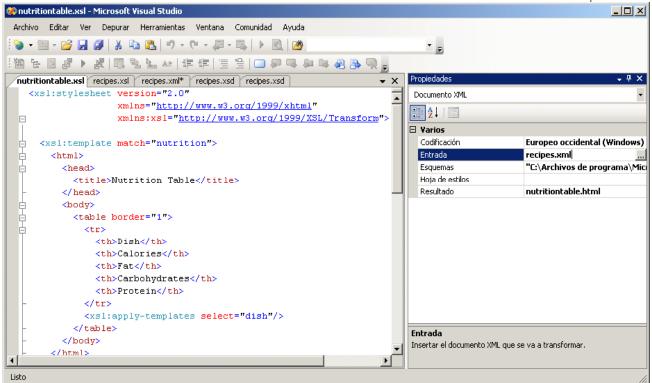


Eskemen editorea (III)

- Definitutako formazioa zuzena den frogatzen du
- Egiaztapena frogatzen du
- Eskemen formatuan nabigazioa eta bistaratzeak (grafikoki)

77

Transformazioen editorea



78



- http://www.w3c.org/Amaya/
 - Amaya. W3C-ko editore/arakatzaile ofiziala
- http://www.vervet.com
 - XML Pro
- http://xmlwriter.net/
 - XMLWriter. XML, DTD, XSL, CSS,... editorea
- http://www.xmlspy.com/
 - XML Spy. XMLeko erabateko editorea
- http://www.oxygenxml.com/index.html
 - <oXygen/> XML Editor & XSLT Debugger
- http://www.stylusstudio.com/
 - Stylus Studio. XMLeko garapen ingurune integratua
- http://tidy.sourceforge.net
 - Tidy. HTML eta XHTML ko kode zuzentzailea
- www.xml.org, www.xml.com,...

79

Aplikazioak

- XML bidez edozein motako dokumentua sor liteke eta bezero aldean datuak maneiatu, zerbitzari aldean motor berezirik behar gabe:
 - Informazioa bilatzen duten aplikazioak
 - Informazio berdina modu desberdinetan bistaratu behar duten aplikazioak (adib., eskuliburuak hizkuntza desberdinetan)
 - Datu-base banatuetan dauden datu-integratuen atzipena behar duten aplikazioak (nahiz haiek formatu desberdinetan egon
 - Bezeroaren aldera karga transferitu behar duten aplikazioak



- XML erabiliz definitu diren lengoaiak kategoria desberdine tan antola litezke:
 - Datuetara zuzendutako lengoaiak: Datu-base erlazionaleko sistematan meta litezkeen datuak deskribatze Datuetara zuzendutako lengoaiak ko erabiliak. Eskemak "lauak" dira (zuhaitzek ez dute sakonera handirik) baina dokumentuak handiak dira ("zabalak", erregistro ugari baitute)
 - Dokumentuetara zuzendutako lengoaiak: XHTML adibidez
 - Protokolo eta prozesaketa lengoaiak: Adibidez WSDL, XSLT edota WML (Wireless markup language) edo ebXML (electronic business)
 - Hibridoak. Aurrekoen ezaugarriak konbinatzen dituzten aplikazioak

81

PHP erabiliz XML prozesatu

- XAMPeko paketeak PHP 5.x instalatu dugu eta harekin:
 - DOM/XML-rekin lan egin dezakegu (enable)
 - Beste aukera bat bertsio murriztuago eta errazago bat erabiltzea da: SimpleXML (luzapen hau ere eskuragarri dago)

XML taldeen datuekin

```
<?xml version="1.0" ?>
<grupos>
<grupo>
  <numero>11</numero>
  <responsable>jperez001@irakasle.ehu.es</responsable>
  <url>http://midominio.es</url>
  <alumnos>
   <alumno>
     <nombre>Jon Pérez</nombre>
     <correo>jperez001@irakasle.ehu.es</correo>
     <titulacion>GI</titulacion>
    </alumno>
    <alumno>
     <nombre>Unai López</nombre>
     <correo>ulopez001@irakasle.ehu.es
      <titulacion>ITIS</titulacion>
    </alumno>
  </alumnos>
 </grupo>
 <grupo>
    ... otro grupo ....
 </grupo>
 .......
</grupos>
```

83

Edukia bistaratu DOMDocument bidez



```
<?php
$doc = new DOMDocument();
$doc->load('grupos.xml');
$root = $doc->documentElement;
$grupos = $doc->getElementsByTagName('grupo');
foreach ($grupos as $child)
{
       $child=$child->firstChild;
              while ($child)
       { if ($child->nodeType==1)
                               echo utf8_decode("$child->nodeName: $child->nodeValue </br>");
         $child = $child->nextSibling;
                                                               http://localhost/ejemploxml/LeerConDOMDocument.php - Windows Internet Explorer
                                                               G ← http://localhost/ejemp ✓ → X S Google
?>
                                                               Archivo Edición Ver Eavoritos Herramientas Ayuda
                                                               Favoritos // http://localhost/ejemploxml/LeerConDOMDocument.php
                                                               numero: 11
                                                                responsable: jperez001@irakasle.ehu.es
                                                                url: http://midominio.es
                                                               alumnos : Jon Pérez j<br/>perez001@irakasle.ehu es GI Unai López ulopez<br/>001@irakasle.ehu.es ITIS numero : 12
                                                                responsable : jsantos001@irakasle.ehu.es
                                                                alumnos : Xabier Alonso xalonso001@irakasle.ehu.es II Julen Santos isantos001@irakasle.ehu.es GI
                                                                                                      🧐 Intranet local

← ● 100% ▼
```





```
<?php
$xml = simplexml_load_file("grupos.xml");
foreach ($xml->as $grupo){
foreach($grupo->children() as $child)
 echo $child->getName() . ": " . $child . "<br />";
 if ($child->getName()=='alumnos'){
     foreach($child->children() as $child2)
                                                                                  🏉 http://localhost/ejemploxml/L... 📳 🗖
                                                                                 C ⇒ E http://localhost/ejemp ▼ ← X
     echo utf8_decode("----". $child2->nombre . "<br/>);
                                                                                           ## http://localhost/ejemploxml/Leer.
 }
                                                                                  responsable: jperez001@irakasle.ehu.es
                                                                                  url: http://midominio.es
                                                                                  alumnos
                                                                                   ----Jon Pérez
?>
                                                                                  -----Unai López
                                                                                  numero: 12
                                                                                  responsable: jsantos001@irakasle.ehu.es
                                                                                  url: http://midominio2.es
                                                                                  -----Xabier Alonso
                                                                                  -----Julen Santos
                                                                                                                  85
                                                                                   Intranet local

← 100%
```

SimpleXML



```
$xml = simplexml_load_file('peliculas.xml');
$xml = simplexml_load_string(String);
```

- Fitxategia prozesatu (aldatu/kontsultatu)
- 3. Gorde

```
$xml->asXML('pelis2.xml');
```

4. Pantailan bistaratu

```
echo $xml->asXML();
```

pelis2.xml

```
pelis2.xml
← → C 🐧 🗋 file:///G:/xampp/htdocs/xml/pelis2.xml
                                                                          ₩.
This XML file does not appear to have any style information associated with it. The document tree
is shown below.
▼<peliculas>
 ▼<pelicula>
    <titulo>La vida de los otros</titulo>
   ▼<personajes>
     ▼<personaje>
        <nombre>Capitán Gerd Wiesler</nombre>
        <actor>Ulrich Mühe</actor>
      </personaje>
     ▼<personaje>
       <nombre>Christa-Maria Sieland</nombre>
        <actor>Martina Gedeck</actor>
      </personaje>
    </personajes>
   ▼<argumento>
      Un equipo de hombres de la Stasi cablea la vivienda de Dreyman,
      instalando micrófonos ocultos, y se dispone la central técnica en el
      desván del edificio
    </argumento>
   ▼<grandes-lineas>
      <linea>Oscar mejor película extranjera 2007</linea>
    </grandes-lineas>
    <puntuacion tipo="pulgares">7</puntuacion>
    <puntuacion tipo="estrellas">5</puntuacion>
   </pelicula>
 ▼<pelicula>
    <titulo>Milenium I</titulo>
   ▼<argumento>
      Un rico industrial llamado Henrik Vanger (Sven-Bertil Taube) contrata
      a un periodista, Mikael Blomqvist (Michael Nyqvist), para saber que
      fue de su sobrina, desaparecida hace nada menos que 40 anios
```



87

Fitxategiaren edukia atzitu



```
//Elementu konkretu bat atzitzen

</php

$xml = simplexml_load_file('pelis2.xml');

$xml->pelicula[1]->argumento='SIN ARGUMENTO';

$xml->asXML('pelis2.xml'); //aldaketak gordetzen ditugu

echo $xml->asXML(); // fitxategia web-arakatzailean bistaratzen da
?>
```





//Ume-bakar ez diren elementuak korritzen:

```
<?php
$xml = simplexml_load_file('pelis2.xml');
echo $xml->pelicula[0]->titulo, '<br>';
// lehenengo pelikularen <puntuacion> nodoak atzitzen ditu
foreach ($xml->pelicula[0]->puntuacion as $puntuacion) {
  switch ((string) $puntuacion['tipo']) { // 'tipo' atributua lortzen du eta haren balioaren idazten du
  case 'pulgares':
     echo $puntuacion, ' pulgares <br/> <br/>'; // hatz lodiak gora
     break:
  case 'estrellas':
     echo $puntuacion, 'estrellas';
                                         // izarrak
     break;
  }
}
?>
```

Fitxategiaren edukia atzitu



```
// Xpath erabili elementuen korritzea

<?php

$pelis = simplexml_load_file('pelis2.xml');

foreach ($pelis->xpath('//pelicula') as $pelicula)

{
   echo utf8_decode("<b>$pelicula->titulo:</b><br>");
   echo "Argumento:<br>";
   echo utf8_decode("$pelicula->argumento<br><br>");
   echo "Actores:<br>";
   foreach ($pelicula->personajes->personaje as $personaje)
    {
      echo utf8_decode("$personaje->nombre - ");
      echo utf8_decode("$personaje->actor<br>");}
   echo "<br/>br>";
}
echo "<br/>cho "<br/>cho">
```



```
<?php
$xml = simplexml_load_file('peliculas.xml');
$pelicula = $xml->addChild('pelicula');
$pelicula->addChild('titulo',$_POST['titulo']); // titulua formulario batetik lortzen dela suposatzen da
$pelicula->addChild('argumento', 'Un rico industrial llamado Henrik Vanger (Sven-Bertil Taube) contrata a un
periodista, Mikael Blomqvist (Michael Nyqvist), para saber que fue de su sobrina, desaparecida hace nada menos
que 40 anios');
$personajes = $pelicula->addChild('personajes');
$personaje - $personajes->addChild('personaje');
$personaje->addChild('nombre', 'Mikael Blomkvist');
$personaje->addChild('actor', 'Michael Nyqvist ');
$puntuacion = $pelicula->addChild('puntuacion', $_POST['puntos']);
$puntuacion->addAttribute('tipo', 'estrellas');
echo $xml->asXML();
$xml->asXML('pelis2.xml');
?>
```

91

SimpleXML eta DOM arteko elkarrekintza



```
<?php
$dom = new domDocument;
$dom->loadXML('<book>><book><title>blah</title></book></book>>');
if (!$dom) {
    echo 'Errorea dokumentua aztertzerakoan (when parsing)';
    exit;
}

$s = simplexml_import_dom($dom);
echo $s->book[0]->title;
?>
```

Beste iturri praktiko batzuk



- how to start simplexml
 - https://www.youtube.com/watch?v=PopVWSBZI6s
- PHP and SimpleXML
 - https://www.youtube.com/watch?v=9xBarPfKvgY
- simpleXML PHP tutorial learn to parse any XML files and RSS feeds
 - https://www.youtube.com/watch?v=4ZLZkdiKGE0