• • • 3.1. HTML

Iturria: Xabier Arregi

Informatika Fakultatea, UPV/EHU

• • HTML

- Helburuak
 - HTMLren ezaugarri oinarrizkoak erakustea
 - HTML dokumentuen egitura aurkeztea
- Aztergaiak
 - Zer da HTML
 - Bilakaera
 - HTML dokumentuen egitura
 - Oinarrizko kontzeptuak
 - HTML elementuak
 - HTML5

• • Zer da HTML

- HyperText Markup Language
- Hipertestuko dokumentu garraiagarriak deskribatzeko lengoaia
- SGMLn oinarritua
- Arakatzaileek bistaratzen dituzte HTML dokumentuak
- Ez da programazio-lengoaia

• • Oinarria: SGML

- Standard Generalized Markup Language
 - Standard: Informazio-trukerako estandarra
 - Generalized: Ez dago aplikazio edo lengoaia bati lotua, zabala da
 - Markup: Markak erabiltzen ditu dokumentuen edukiak deskribatzeko
 - Language: Dokumentuen egitura azaleratzeko sintaxia finkatzen du

• • Oinarria: SGML (II)

- SGMLren osagaiak:
 - DTD: Dokumentu-motaren definizioa
 - Erabiliko diren markak finkatzen ditu
 - Marka horiek noiz eta nola jarri behar diren erakusten du
 - Markei atributuak egokitzen dizkie
 - Dokumentu instantzia
 - DTDan definitutako markak eta entitateak dauzka
 - Dokumentua marka horiekin etiketatuta dago
- HTML, hortaz, hipertestuak deskribatzeko DTDa da

W3C Recommendation

previous next contents elements attributes index

21 Document Type Definition

```
<!--
    This is HTML 4.01 Strict DTD, which excludes the presentation
    attributes and elements that W3C expects to phase out as
    support for style sheets matures. Authors should use the Strict
    DTD when possible, but may use the Transitional DTD when support
    for presentation attribute and elements is required.
    HTML 4 includes mechanisms for style sheets, scripting,
    embedding objects, improved support for right to left and mixed
    direction text, and enhancements to forms for improved
    accessibility for people with disabilities.
          Draft: $Date: 1999/12/24 23:37:49 $
          Authors:
              Dave Raggett <dsr@w3.org>
              Arnaud Le Hors <lehors@w3.org>
              Ian Jacobs <ij@w3.org>
    Further information about HTML 4.01 is available at:
        http://www.w3.org/TR/1999/REC-html401-19991224
    The HTML 4.01 specification includes additional
    syntactic constraints that cannot be expressed within
    the DTDs.
-->
<!--
    Typical usage:
    <!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN"
            "http://www.w3.org/TR/html4/strict.dtd">
    <html>
    <head>
    </head>
```

W3C Recommendation

<!--HTML has two basic content models: %inline; character level elements and text strings %block; block-like elements e.g. paragraphs and lists --> <!ENTITY % block "P | %heading; | %list; | %preformatted; | DL | DIV | NOSCRIPT | BLOCKQUOTE | FORM | HR | TABLE | FIELDSET | ADDRESS"> <!ENTITY % flow "%block; | %inline;"> <!--============ Document Body ==================-----<!ELEMENT BODY 0 0 (%block; |SCRIPT)+ +(INS|DEL) -- document body --> <!ATTLIST BODY %attrs; -- %coreattrs, %i18n, %events --#IMPLIED -- the document has been loaded -onload %Script; #IMPLIED -- the document has been removed -onunload %Script; <!ELEMENT ADDRESS - - (%inline;)* -- information on author --> <!ATTLIST ADDRESS -- %coreattrs, %i18n, %events --%attrs; <!ELEMENT <u>DIV</u> - - (<u>%flow;</u>)* -- generic language/style container --> <!ATTLIST DIV %attrs; -- %coreattrs, %i18n, %events ---- reserved for possible future use --%reserved; <!ENTITY % Shape "(rect|circle|poly|default)"> <!ENTITY % Coords "CDATA" -- comma-separated list of lengths --> <!ELEMENT A - - (%inline;)* -(A) -- anchor --> <!ATTLIST A %attrs; -- %coreattrs, %i18n, %events --#IMPLIED -- char encoding of linked resource -charset %Charset;

• • HTML: Historia eta bilakaera

- Tim Berners-Leek (webaren asmatzaileak) sortua 1991n
- World Wide Web Consortium (W3C) erakundeak kontrolatzen du
 - http://www.w3.org/
- HTML 2.0 (1995)
 - Estandarra oinarrizko ezaugarrietarako.
- HTML 3.2 (1997)
 - Ezaugarri garrantzitsuak gehitzen ditu: taulak, applet-ak, irudiak, etab. Proposamena.
- HTML 4.0 (1997-98)
 - Eleaniztasuna
 - Taulak, formularioak eta script-ak tratatzeko hobekuntzak.
 - Markoen eta estilo-orrien erabilera: aurkezpena eta egitura bereizteko saioa
 - HTML 4.01 (1999ko abendua) azkeneko bertsio onartua da: sintaxi definizio zehatza eta formala

• • HTML. Bilakaera (II)

- HTMLri buruzko gomendioen geldialdia
- XHTML (Extensible HyperText Markup Language)
 - HTML XMLz birformulatzea:
 - XHTML 1: 1999n, 2000n estandarra
 - XHTML 1.0 = HTML 4.01 XMLz inplementatua
 - Modularizazioa: XHTML 1.1 (2001eko maiatza)

- 2004an WHATWG taldea sortu zen
 - Web Hypertext Application
 Technology Working Group
 - Apple, Mozilla, Opera, Google...
 - HTML 5-en lanean hasi ziren

HTML dokumentuen egitura

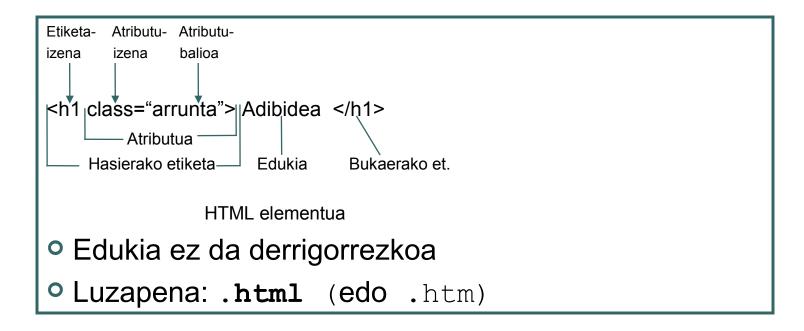
 Elementuak adierazteko hasierako eta bukaerako markak (tag) erabiltzen dira:
 <etiketa> </etiketa>

- Erabil daitezkeen marken multzoa aurrez definituta dago
- Markek atributuak eduki ditzakete

<etiketa atributu1...> </etiketa>

<etiketa atributu1=balio1...> </etiketa>

HTML dokumentuen egitura (II)



HTML dokumentuen egitura (III)

<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0//EN"

"http://www.w3.org/TR/REC-html40/strict.dtd"

```
<html>
   <head>
       <title> Izenburua </title>
       Goiburuko beste zenbait elementu
   </head>
   <body>
      Dokumentuaren gorputza
   </body>
</html>
```

17

Dokumentu-motaren erazagupena.

• • Goiburukoa: HEAD

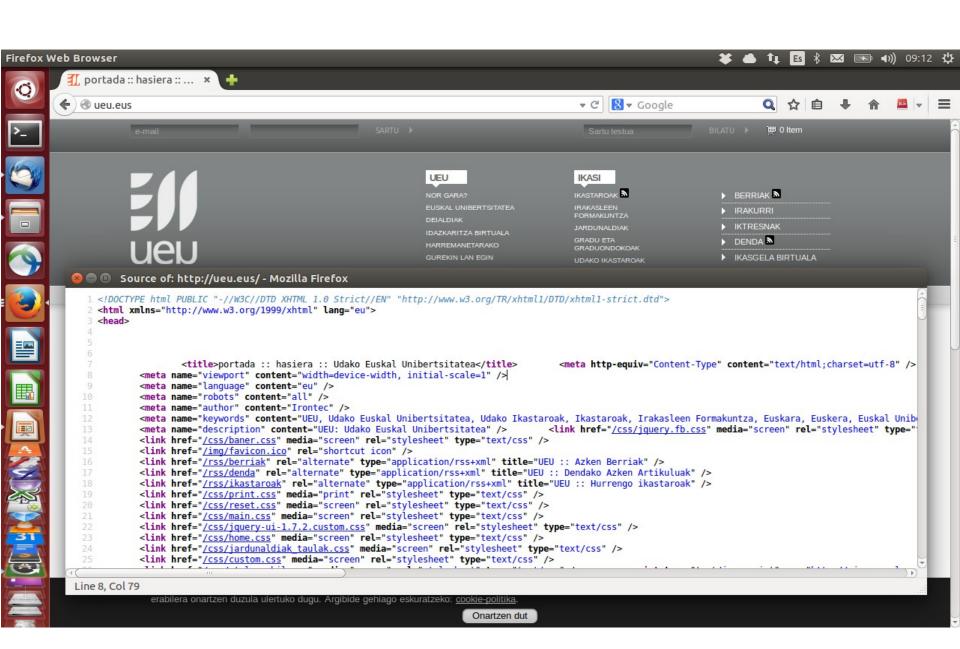
 Dokumentuari buruzko informazio orokorra Adibidea:

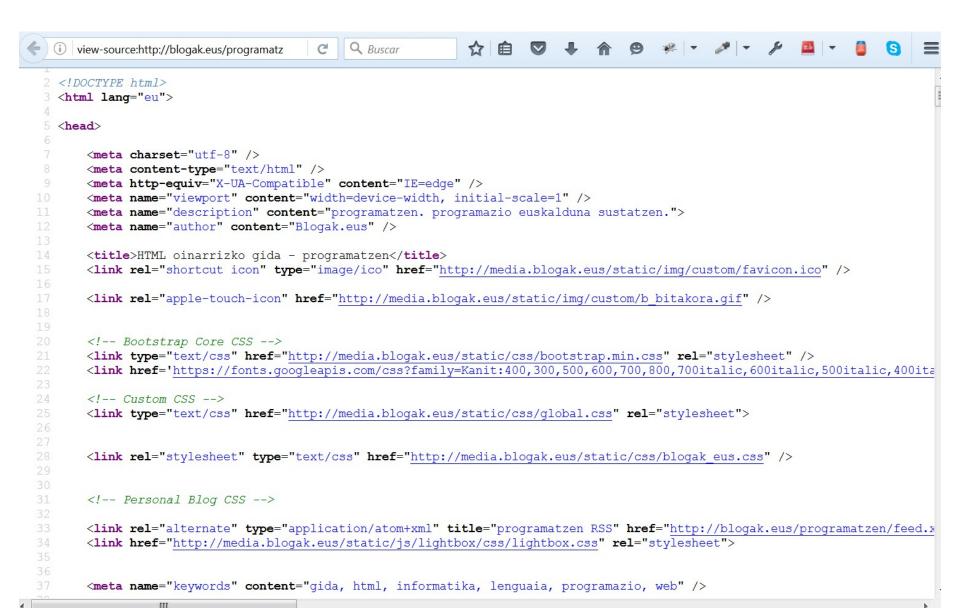
```
<head>
     <meta http-equiv="Content-Type" content="text/html;
     charset=UTF-8" />
        <title>SGTA</title>
</head>
```

- Ez da zuzenean bistaratzen
- Azpiprogramak goiburukoan definitzen dira (JavaScript, VBScript)

• • Goiburuko elementuak

- <base /> Oinarriko URLa. Erreferentzia partzialak ebazteko.
- <meta /> Metainformazioa: hitz gakoak, deskribapena, egilea.
- <title> Dokumentuaren izenburua definitzen du. Beti da beharrezkoa.
- link /> Dokumentuen arteko erlazioa.
 Baliabide garrantzitsu batekiko esteka adierazten da.
- <style> Estilo edukiak adierazteko.
- <script> Bezeroko script bat.





• • Gorputza: BODY

- Testua jartzen da, irakurleak ikusiko dituen efektuak eta estekak barne.
- Gorputzean edukiko dira:
 - Gorputzaren egiturazko elementuak
 - Hiperestekak
 - Testua markatzeko elementuak
 - Listak
 - Taulak
 - Irudiak eta mapak
 - Objektuak
 - Formularioak

Gorputzaren egiturazko elementuak

- Iruzkinak: <!-- iruzkina, lerro anitz -->
- <h[1..6]> goiburukoak (lehen mailatik seigarrenera)
- <div> edukiontzi generikoa
- paragrafo baten edo lerro baten baitako edukiontzia
- paragrafoa
-
 lerro-etena
- <hr /> marra horizontala

• • Hiperestekak

- Hiperestekak definitzeko
 - href atributua: helburuaren URLa

```
<a href="http://www.jalgi.com"> Testua </a>
```

 Testua

 Testua

• Aingurak:

```
<a name="behera"> </a>
```

. . .

 Testua

Testua markatzeko elementuak

- enfasia, letra lodia
- enfasia, letra etzana
- <q> aipu laburra
- <abbr> laburdura. Forma hedatua title atributuan adierazten da
- <acronym> akronimoa.
 Forma hedatua title
 atributuan adierazten da
- <address> harremanetarako informazioa

- testu aurreformateatua
- <dfn> termino bat, definiturik agertzen dela
- <code> programa baten kodea
- <cite> izenburuen aipuak...
- ezabatu den edukia
- <ins> txertatu den edukia
- <sub> azpi-indizea
- <sup> goi-indizea
- <bd><bdo> testuaren noranzkoa

• • Listak

- Elementuen segida bat aurkezteko erabiltzen dira listak
- Hiru mota:
 - ordenatuak:
 - desordenatuak:
 - definizio-listak: <dl>
- Lista bat definitzeko adierazi behar da:
 - Lista non hasten den: ,
 - Elementu bakoitza non hasi eta non bukatzen den. (,)
 - Lista non bukatzen den. (,)
- O Definizio-listak: (<dl>,</dl>)
 - Definizio-terminoak: (<dt>,</dt>)
 - Deskribapenak: (<dd>,</dd>)

• • Taulak

- taula
- <caption> izenburua
- Egitura:
 - Goiburukoa: <thead>
 - Gorputza:
 - Oina: <tfoot>
- Taulako errenkadak:
- Gelaskak:

 - goiburukoetan.
- Zutabeak biltzeko: <colgroup>

• • Irudiak eta mapak

- irudiak
- <map id="bertxin" name="xxx">
 - id edo name atributuak mapa identifikatzeko izenak dira, aingura moduan erabiltzen direnak. Mapan kokatuko den objektuak edo irudiak usemap atributuan erabiliko ditu.
- <area alt="" /> mapa bateko area

• • Objektuak

- <object> objektua: irudia, audioa, bideoa, Java appleta, Flash animazioa.
- o <param name="" /> objektu baten parametroak adierazteko (applet-etan)

• • Formularioak

- Formulario interaktiboa: Erabiltzaileari datuak idazteko aukera emateko.
 - Bere baitan kontrolak biltzen ditu: testu-kutxak, testu-areak, botoiak...

• • Karaktere maila

- Zuriune-kateak zuriune bakarraren parekoak dira
- Karaktere-entitateak: zenbait karaktere modu berezian idazten da:

```
& eta; letren artean. Adb:
```

```
< <
&gt; >
&amp; &
&quot; 
&shy; -
&aacute; &
&ntilde; f
```

• • Atributuak.

- Funtsezkoak:
 - class: estiloa adierazteko
 - id: identifikazioa. Indize modukoa.
- Hizkuntzari buruzkoak:
 - lang: hizkuntza
 - dir: noranzkoa

• • HTMLren gabeziak

- Ez da zabalgarria
- Aurkezpeneko aspektuetan zentratua
- Edukia aurkezpenarekin nahasten du
- Ez du egitura semantikorik
- Ondorioa:

HTML ez da egokia datuak errepresentatzeko

- Hertsiegia delakoan...
 - XML: eXtensible Markup Language
 - SGMLren antzekoa, baina hipertestuetarako

• • • XHTML

- HTMLren aldaera bat baizik ez da
 - Sintaxiaren aldetik, XHTML zorrotzagoa eta erregularragoa da.
- Arauak:
 - Etiketak eta atributuak minuskulaz
 - Bukaerako etiketak derrigorrezkoak ()
 - Ixtekorik behar ez duten etiketek "/" karakterea eraman behar dute (
)

• • XHTML: errore tipikoak

- Elementu hutsak ez ondo ixtea
 - Okerra:

 - Zuzena:

- Elementu ez-hutsak ez ondo ixtea
 - Okerra: This is a paragraph.This is another paragraph.
 - Zuzena: This is a paragraph.This is another paragraph.
- Elementuak ez ondo habiatzea
 - Okerra: This is some text.
 - Zuzena: This is some text.
- Komatxoak ez jartzea atributuen balioetan
 - Okerra:
 - Zuzena:
 - Zuzena:
- Ampersand karakterea entitate batetik kanpo erabiltzea
 - Okerra: <title>Cars & Trucks</title>
 - Zuzena: <title>Cars & Trucks</title>

XHTML: errore tipikoak (II)

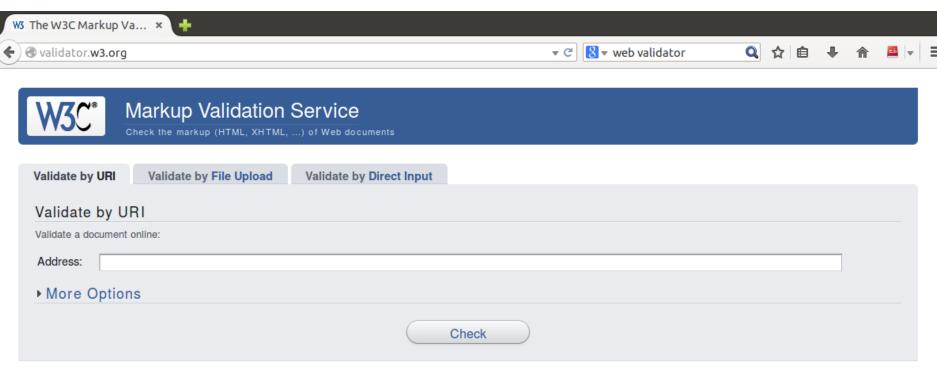
- URLetan, ampersand karakterea entitate gisa ez erabiltzea
 - Okerra: News
 - Zuzena: News
- Etiketetan eta atributuetan, maiuskulak eta minuskulak ez bereiztea
 - Okerra: <BODY><P ID="ONE">The Best Page Ever</P></BODY>
 - Zuzena: <body>The Best Page Ever</body>
- Atributuen minimizazioa erabiltzea
 - Okerra: <textarea readonly>READ-ONLY</textarea>
 - Zuzena: <textarea readonly="readonly">READ-ONLY</textarea>
- Karaktere-kodeketa ez adieraztea.
 - <meta http-equiv="Content-Type" content="text/html; charset=utf-8" />
- HTML elementuaren izen-espazioa ez jartzea:
 - <html xmlns="http://www.w3.org/1999/xhtml">
- Root elementua:
 - <?xml version="1.0" encoding="UTF-8"?>

• • XHTML: bilakaera

- XHTML 1.0 (1997)
 - lehenengo bertsioa.
 - HTML 4.01-en birformulazioa XMLn.
 - Strict, Transitional, Frameset
- XHTML Basic
 - Oinarrizko ezaugarriak, bezero sinpleentzat
- Modularization of XHTML
 - Etiketak moduluka antolatzen dira
 - Posible da konbinatzea modulu (DTD) desberdinak: vector graphics, multimedia, math, electronic commerce...
- XHTML 1.1 Module-based XHTML
 - Moduluak integratuz osatzen dira dokumentu motak. Adb: XHTML 1.1 plus MathML 2.0
- XHTML 2.0: lanean ari dira

• • Balidazioa

- Egokitzapen automatikoa egiteko tresna:
 - HTML Tidy: http:// tidy.sourceforge.net/, http://cgi.w3.org/cgi-bin/tidy
- Egiaztatzeko tresna:
 - http://validator.w3.org/



This validator checks the markup validity of Web documents in HTML, XHTML, SMIL, MathML, etc. If you wish to validate specific content such as RSS/Atom feeds or CSS stylesheets, MobileOK content, or to find broken links, there are other validators and tools available. As an alternative you can also try our non-DTD-based validator.



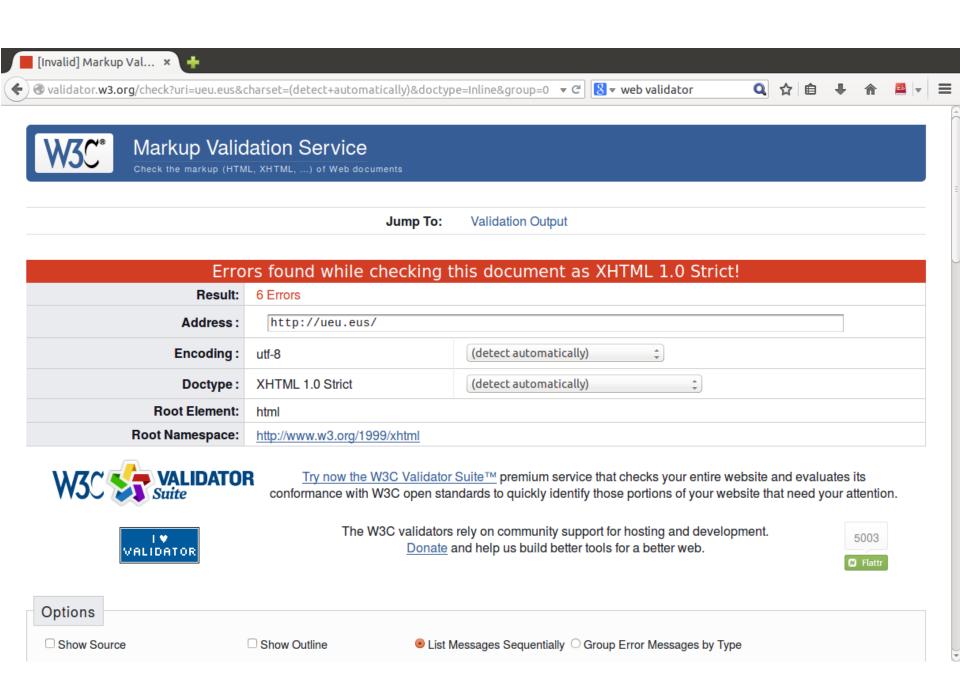
<u>Try now the W3C Validator Suite™</u> premium service that checks your entire website and evaluates its conformance with W3C open standards to quickly identify those portions of your website that need your attention.

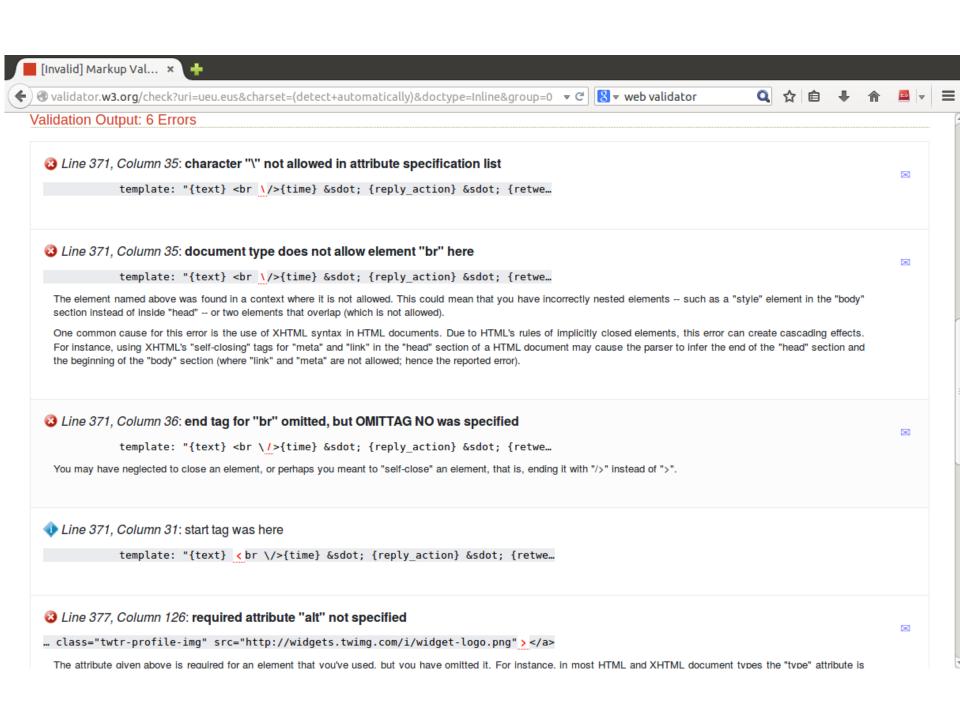


The W3C validators are hosted on server technology donated by HP, and supported by community donations.

<u>Donate</u> and help us build better tools for a better web.







Web orrien diseinuan izandako bilakaera

- Aldaketa sakona orrien diseinuan.
- Lehen:
 - HTML 2.0, HTML 3.2: bisualizazioko informazioa elementuen baitan.
 - HTML etiketen eta atributuen bidez lantzen zen formatua.
- Orain:
 - Hiru eginkizun nahiko independente:
 - Orriaren eduki eta egitura logikoa (HTML)
 - Formatu-etiketarik ez litzateke erabili beharko (adb.)
 - Orriaren estiloa (CSS)
 - Orriaren aspektu dinamikoak (DOM + script-ak)

• • Helbide interesgarriak

HTML

- http://www.w3.org/MarkUp/
- http://validator.w3.org/
 - **HTML Validation Service**
- http://www.w3.org/TR/html401/
 HTML 4.01 lengoaiaren espezifikazioa
- http://www.htmlhelp.com/reference/html40/
 HTML 4.0 lengoaiaren erreferentzia
- http://www.w3.org/TR/html401/struct/global.html
 HTML 4.01 dokumentuen egitura
- http://en.wikipedia.org/wiki/Markup_language
- Tutoriala: http://www.w3schools.com/tags/default.asp

- Tutoriala euskaraz: http://www.blogak.com/programatzen/html-oinarrizko-gida
- Txuleta:
 http://www.cril.univ-artois.fr/~lecoutre/teaching/web/sheets/HTML.pdf

XHTML

- Euskarazko wikipedia: http://eu.wikipedia.org/wiki/XHTML
- Ingelesezko wikipedia: http://en.wikipedia.org/wiki/XHTML
- W3C-ren espezifikazioa: http://www.w3.org/TR/xhtml11/
- Txuleta: http://floele.flyspray.org/htmlcheatsheet.pdf