

G3.1(a). HTML

Iturria: Xabier Arregi

Informatika Fakultatea, UPV/EHU



Zer da HTML

- HyperText Markup Language
- Hipertestuko dokumentu garraiagarriak **deskribatzeko** lengoaia
- SGMLn oinarritua
- Arakatzailleek 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



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

<http://www.evolutionoftheweb.com/?hl=es>



HTML dokumentuen egitura

- **Elementuak** adierazteko hasierako eta bukaerako markak (tag) erabiltzen dira:

<etiketa> </etiketa>

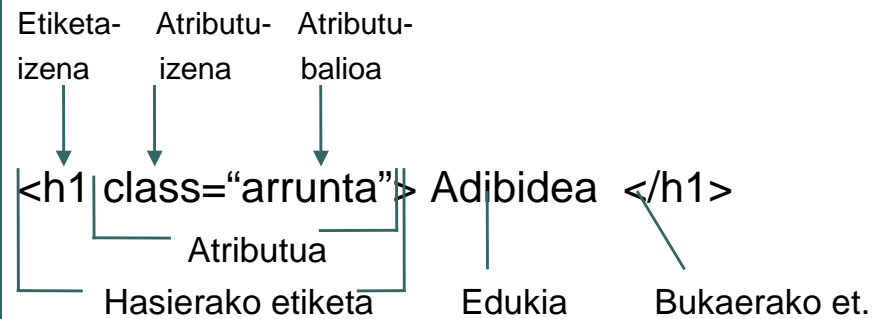
- Erabil daitezkeen **marken multzoa** aurrez definituta dago

- **Markek atributuak** eduki ditzakete

<etiketa atributu1...> </etiketa>

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

HTML dokumentuen egitura (II)



HTML elementua

- Edukia ez da derrigorrezkoa
- Luzapena: **.html** (edo `.htm`)

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>
```

*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>` - Bezero aldeko script bat.


Firefox Web Browser

portada :: hasiera :: ... x

ueu.eus

Google

e-mail SARTU Sartu testua BILATU 0 Item

 **UEU**
NOR GARA?
EUSKAL UNIBERTSITATEA
DEIALDIK
IDAZKARITZA BIRTUALA
HARREMANETARAKO
GUREKIN LAN EGIN

IKASI
IKASTAROAK
IRAKASLEEN
FORMAKUNTZA
JARDUNALDIK
GRADU ETA
GRADUONDOKOAK
UDAKO IKASTAROAK

BERRIAK
IRAKURRI
IKTRESNAK
DENDA
IKASGELA BIRTUALA

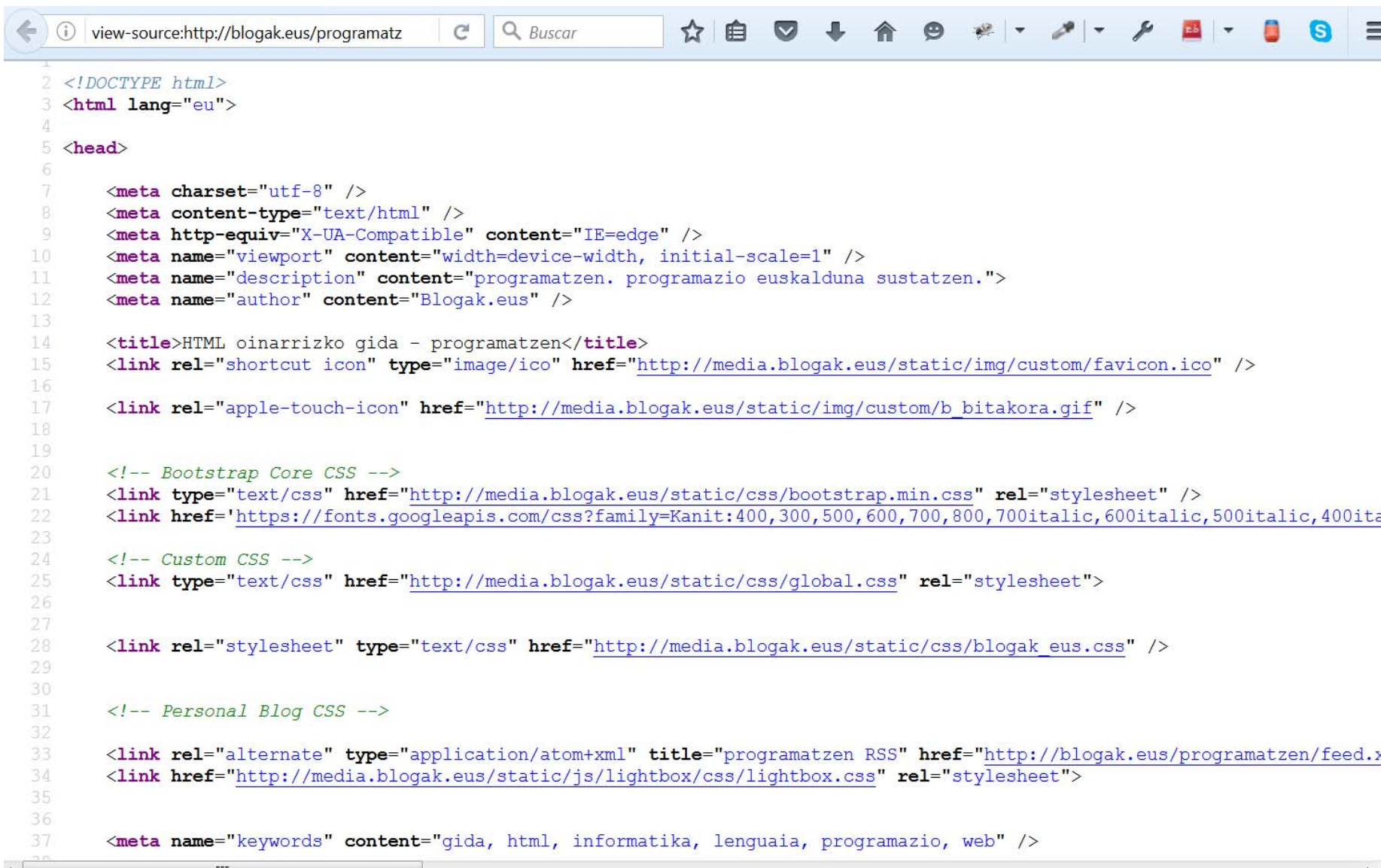
Source of: http://ueu.eus/ - Mozilla Firefox

```
1 <!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
2 <html xmlns="http://www.w3.org/1999/xhtml" lang="eu">
3 <head>
4
5
6
7     <title>portada :: hasiera :: Udako Euskal Unibertsitatea</title>
8     <meta http-equiv="Content-Type" content="text/html; charset=utf-8" />
9     <meta name="viewport" content="width=device-width, initial-scale=1" />
10    <meta name="language" content="eu" />
11    <meta name="robots" content="all" />
12    <meta name="author" content="Irontec" />
13    <meta name="keywords" content="UEU, Udako Euskal Unibertsitatea, Udako Ikastaroak, Ikastaroak, Irakasleen Formakuntza, Euskara, Euskera, Euskal Unib" />
14    <meta name="description" content="UEU: Udako Euskal Unibertsitatea" />
15    <link href="/css/jquery.fb.css" media="screen" rel="stylesheet" type="text/css" />
16    <link href="/css/baner.css" media="screen" rel="stylesheet" type="text/css" />
17    <link href="/img/favicon.ico" rel="shortcut icon" />
18    <link href="/rss/berriak" rel="alternate" type="application/rss+xml" title="UEU :: Azken Berriak" />
19    <link href="/rss/denda" rel="alternate" type="application/rss+xml" title="UEU :: Dendako Azken Artikuluak" />
20    <link href="/rss/ikastaroak" rel="alternate" type="application/rss+xml" title="UEU :: Hurrengo ikastaroak" />
21    <link href="/css/print.css" media="print" rel="stylesheet" type="text/css" />
22    <link href="/css/reset.css" media="screen" rel="stylesheet" type="text/css" />
23    <link href="/css/main.css" media="screen" rel="stylesheet" type="text/css" />
24    <link href="/css/jquery-ui-1.7.2.custom.css" media="screen" rel="stylesheet" type="text/css" />
25    <link href="/css/home.css" media="screen" rel="stylesheet" type="text/css" />
26    <link href="/css/jardunaldiak_taulak.css" media="screen" rel="stylesheet" type="text/css" />
27    <link href="/css/custom.css" media="screen" rel="stylesheet" type="text/css" />
28 </head>
29 <body>
30
31 </body>
32 </html>
```

Line 8, Col 79

erabilera onartzen duzula ulertuko dugu. Argibide gehiago eskuratzeko: [cookie-politika](#).

Onartzen dut





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
- `<p>` - paragrafoa
- `
` - lerro-etena
- `<hr />` - marra horizontala



Hiperestekak

- `<a>` Hiperestekak definitzeko (anchor)
- `` href atributua: helburuaren URLa
 - ` Testua `
 - ` Testua `
 - ` John Doe `
- Aingurak:
 - ` `
 - ...
 - ` Testua `



Testua markatzeko elementuak

- **** - enfasia, letra lodia
- **** - enfasia, letra etzana
- **<blockquote>** - bloke mailako aipua
- **<q>** - aipu laburra
- **<abbr>** - laburdura. Forma hedatua **title** atributuan adierazten da
- **<acronym>** - akronimoa. Forma hedatua **title** atributuan adierazten da
- **<address>** - harremanetarako informazioa
- <pre>** - 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
- <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. (``, ``)
- Definizio-listak: (`<dl>`, `</dl>`)
 - Definizio-terminoak: (`<dt>`, `</dt>`)
 - Deskribapenak: (`<dd>`, `</dd>`)



Taulak

- **<table>** - taula
- **<caption>** - izenburua
- Egitura:
 - Goiburukoa: **<thead>**
 - Gorputza: **<tbody>**
 - Oina: **<tfoot>**
- Taulako errenkadak: **<tr>**
- Gelaskak:
 - **<td>**
 - **<th>** - 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.
- **<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:

<	<
>	>
&	&
"	"
­	-
á	á
ñ	ñ



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 (`</p>`)
 - 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: `<p>This is a paragraph.<p>This is another paragraph.`
 - Zuzena: `<p>This is a paragraph.</p><p>This is another paragraph.</p>`
- Elementuak ez ondo habiatzea
 - Okerra: `This is some text.`
 - Zuzena: `This is some text.`
- Komatxoak ez jartzea atributuen balioetan
 - Okerra: `<td rowspan=3>`
 - Zuzena: `<td rowspan="3">`
 - Zuzena: `<td rowspan='3'>`
- *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><p id="ONE">The Best Page Ever</p></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/>



Markup Validation Service

Check the markup (HTML, XHTML, ...) of Web documents

Validate by URI

Validate by File Upload

Validate by Direct Input

Validate by URI

Validate a document online:

Address:

► [More Options](#)

Check

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](#).



[Try now the W3C Validator Suite™](#) 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. [Donate](#) and help us build better tools for a better web.

5003

Flattr

W3C[®] Markup Validation Service

Check the markup (HTML, XHTML, ...) of Web documents

Jump To: [Validation Output](#)

Errors found while checking this document as XHTML 1.0 Strict!

Result: 6 Errors

Address :

Encoding : utf-8 (detect automatically)

Doctype : XHTML 1.0 Strict (detect automatically)

Root Element: html

Root Namespace: <http://www.w3.org/1999/xhtml>



[Try now the W3C Validator Suite™](#) 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 rely on community support for hosting and development.
[Donate](#) and help us build better tools for a better web.

5003



Options

☐ Show Source

☐ Show Outline

☒ List Messages Sequentially ☐ Group Error Messages by Type

Validation Output: 6 Errors

✖ Line 371, Column 35: character "\" not allowed in attribute specification list

```
template: "{text} <br \"/>{time} &sdot; {reply_action} &sdot; {retwe...
```

✖ Line 371, Column 35: document type does not allow element "br" here

```
template: "{text} <br \"/>{time} &sdot; {reply_action} &sdot; {retwe...
```

The element named above was found in a context where it is not allowed. This could mean that you have incorrectly nested elements -- such as a "style" element in the "body" section instead of inside "head" -- or two elements that overlap (which is not allowed).

One common cause for this error is the use of XHTML syntax in HTML documents. Due to HTML's rules of implicitly closed elements, this error can create cascading effects. For instance, using XHTML's "self-closing" tags for "meta" and "link" in the "head" section of a HTML document may cause the parser to infer the end of the "head" section and the beginning of the "body" section (where "link" and "meta" are not allowed; hence the reported error).

✖ Line 371, Column 36: end tag for "br" omitted, but OMITTAG NO was specified

```
template: "{text} <br \"/>{time} &sdot; {reply_action} &sdot; {retwe...
```

You may have neglected to close an element, or perhaps you meant to "self-close" an element, that is, ending it with "/" instead of ">".

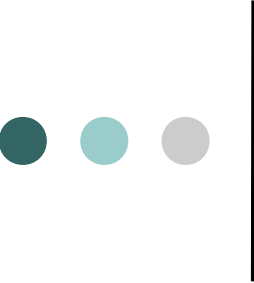
ⓘ Line 371, Column 31: start tag was here

```
template: "{text} <br \"/>{time} &sdot; {reply_action} &sdot; {retwe...
```

✖ Line 377, Column 126: required attribute "alt" not specified

```
... class="twtr-profile-img" src="http://widgets.twimg.com/i/widget-logo.png" > </a>
```

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



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)