

Università Roma Tre

# HTML e XHTML

Questo lavoro è concesso in uso secondo i termini di una licenza Creative Commons  
(vedi ultima pagina)

Sistemi Distribuiti aa 2003-2004

Università Roma Tre

## Sommario

- HTML: semantica di visualizzazione
- Storia di HTML
- Principali problemi tecnologici
- Standard rilevanti
  - HTML 1.0 e CSS
- XHTML
  - convalida del codice
  - uso degli standard

Sistemi Distribuiti aa 2003-2004

Università Roma Tre

## Introduzione

- HTML
  - un linguaggio per descrivere documenti ipermediali
  - documenti fatti di titoli, paragrafi, testo ecc.
  - collegamenti (link) ad altri documenti
  - contenuti di altri media (es: immagini)
- Linguaggio di marcatura
  - obbedisce ad un DTD di SGML

Sistemi Distribuiti aa 2003-2004

Università Roma Tre

## Introduzione

- SGML
  - sintassi per descrivere la struttura gerarchica di documenti
  - il “padre” di XML (anni '70); più complesso
  - DTD: schemi per i documenti
- Meno restrittivo di XML
  - i tag di chiusura possono mancare
  - maiuscole e minuscole
  - apici per gli attributi

Sistemi Distribuiti aa 2003-2004

## Introduzione

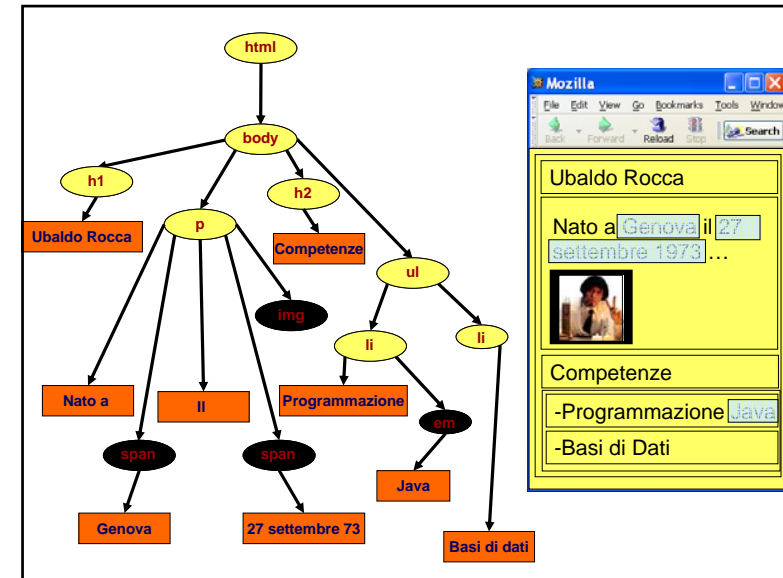
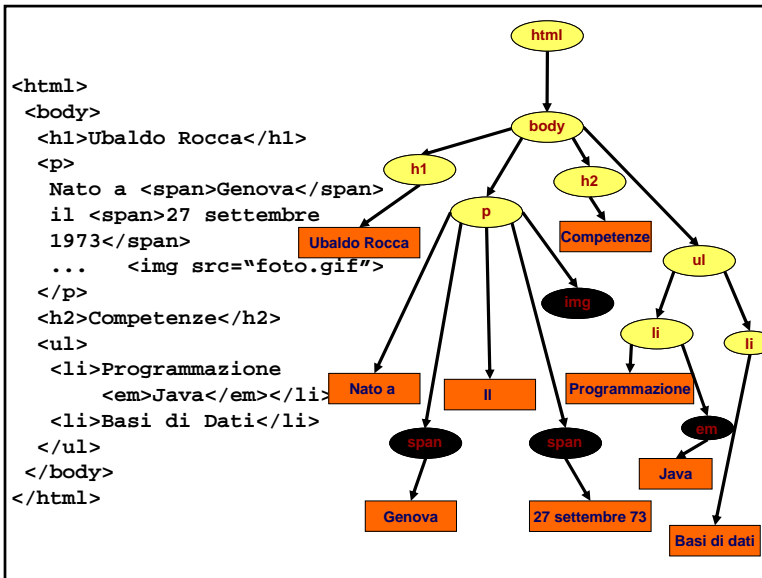
```
<html>
<head>
  <title> Il portale universale </title>
</head>
<body>
  <h1>Portale Universale</h1>
  <p>Questa &grave; la pagina principale del
    <b>Portale Universale</b>.
    Materiale Interessante</a>
    ...
    <br />
    <i>Ultima modifica 23 Febbraio 2004</i>
  </p>
</body>
</html>
```

Sistemi Distribuiti aa 2003-2004

## HTML: semantica di visualizzazione della pagina

- I documenti HTML corrispondono ad alberi
- In fase di visualizzazione, la nidificazione dei nodi corrisponde ad una nidificazione di riquadri
- Distinguiamo riquadri
  - *Inline*: il flusso non viene interrotto
  - *Block*: il flusso viene interrotto

Sistemi Distribuiti aa 2003-2004



## La Storia di HTML

- HTML 1.0: La prima versione di HTML
  - Tim Berners Lee, Dave Raggett, giugno '93
- Il DTD di HTML 1.0
  - titoli ("headings")  
Es. <h1>, <h2>, <h3>, ..., <h6>
  - paragrafi e formattazione del testo  
Es. <p>, <br>, <b>, <i>, <b> ...
  - liste e glossari  
Es. <ul>, <ol>, <li>, ...
  - collegamenti ad altre pagine  
Es. <a>
  - Immagini  
Es. <img>

## La Storia di HTML

- Primi browser testuali
  - Lynx, 1991 – 1992
- Primi browser grafici
  - Mosaic 1.0  
novembre '93, Marc Andreessen, NCSA-UIUC
  - Mosaic 2.0 v. alfa  
gennaio '94, introduce le maschere ("form")
- 1994
  - Andreessen e Jim Clark fondano la Netscape

## La Storia di HTML

- W3C
  - ottobre '94, fondato dal MIT
  - comincia l'attività di standardizzazione
- HTML 2.0 (RFC 1866)
  - tra l'aprile '94 e il settembre 1995
  - standardizza le maschere e le mappe
- Netscape Navigator 1.0
  - dicembre '94, supporto HTML 2.0 e nuovi tag

## La Storia di HTML

- Netscape 1.1
  - aprile '95, introduce le tabelle e nuovi tag
- Internet Explorer 1.0
  - agosto '95, distribuito gratuitamente
- Nel 1995
  - Netscape ha l'**80%** del mercato
  - nascono i primi motori di ricerca
  - inizia la "guerra dei browser" (*browser war*)

## La Storia di HTML

- HTML 3.0 (draft W3C)
  - marzo '95 – settembre '95, tabelle
- Internet Explorer 2.0
  - novembre '95, tabelle <marquee>
- Netscape Navigator 2.0
  - marzo '96, frame, plug-in, applet, Javascript
- CSS1, Cascading Style Sheets Livello 1
  - novembre '95 – dicembre '96

## La Storia di HTML

- HTML 3.2
  - standard W3C, maggio '96 – gennaio '97
  - introduce script e tag di presentazione <font>
- Netscape Navigator 3.0
  - agosto '96, tag <multicol>, <spacer>
- Internet Explorer 3.0
  - agosto '96, JScript, VBScript

## La Storia di HTML

- Netscape Navigator 4.0
  - giugno '97, layer, CSS1
- Internet Explorer 4.0
  - ottobre '97, versione migliorata, CSS1
- Gennaio 1998
  - la quota di Netscape è scesa al **40%** circa
  - il browser diventa gratuito
  - strategia “open source” (nasce Mozilla)

## La Storia di HTML

- Si affacciano altre tecnologie
  - DHTML “Dynamic HTML”
  - combinazione di DOM, CSS e JavaScript
  - possibilità di animare elementi della pagina
- Macromedia Flash
  - grafica vettoriale
  - animazioni
- Alcuni fallimenti autorevoli

## La Storia di HTML

- HTML 4.0
  - giugno '97 – aprile '98
  - introduce i fogli di stile, standardizza i frame
  - DTD “strict”: solo elementi strutturali
  - DTD “transitional”
  - DTD “frameset”
- CSS2
  - maggio '98, posizionamento assoluto

## La Storia di HTML

- Internet Explorer 5.0
  - marzo '99, supporto per XML, CSS1 e CSS2
- Netscape 6
  - novembre 2000, abbastanza instabile
- Ottobre – Novembre 2001
  - Netscape 6.2
  - Internet Explorer 6
  - Opera 6

## La Storia di HTML

- Perché questa storia
  - consente di capire lo stato della tecnologia
  - e i nuovi standard del W3C
- Alla fine degli anni '90
  - mercato stabilizzato
  - il Consorzio finalmente autorevole
  - un pesante retaggio tecnologico
  - ulteriori evoluzioni tecnologiche e sociali (dispositivi mobili, disabili ecc.)

## Principali Problemi del Linguaggio

- Problema 1: Evoluzione molto rapida
  - mancanza di metodologie consolidate
  - struttura e presentazione sono indipendenti
- Indicazione del Consorzio
  - separare struttura e presentazione
  - CSS livello 1 e CSS livello 2

## Principali Problemi del Linguaggio

- Problema 2: Codice scorretto
  - i browser sono “onnivori”
  - i DTD dello standard non sono rispettati
  - problemi per le applicazioni
- Indicazione del Consorzio
  - adottare la sintassi XML
  - XHTML 1.0: una riformulazione di HTML 4.0 secondo lo standard di XML

## Principali Problemi del Linguaggio

- Problema 3: Diverse capacità dei browser
  - diverse versioni di HTML, diversi DTD
  - diverse piattaforme (palmari, WebTV, ...)
- Indicazione del Consorzio
  - modularizzazione di XHTML e negoziazione
  - principio di “accessibilità”
  - estensibilità
  - “Modularization of XHTML”, XHTML1.1

## Standard Rilevanti

- XHTML 1.0
  - gennaio 2000
- CSS livello 1, CSS livello 2
  - dicembre '96, maggio '98
- Modularization of XHTML, XHTML 1.1
  - aprile 2001, maggio 2001
- XHTML Basic
  - dicembre 2000

## XHTML 1.0

- Riformulazione di HTML 4.0
  - formato e DTD XML e non SGML
- In altri termini
  - i documenti XHTML corretti sono alberi validi rispetto ad un DTD XML fissato dal consorzio
  - il DTD descrive quali elementi e quali attributi possono comparire in una pagina XHTML

## XHTML 1.0

- DTD di riferimento
  - DTD strict **XHTML1-strict.dtd**  
è quello di riferimento; presentazione CSS
  - DTD transitional XHTML1-transitional.dtd  
è da intendersi come temporaneo
  - DTD frameset XHTML1-frameset.dtd  
è di fatto “deprecato”
- Ci occuperemo solo di XHTML1-strict

## XHTML 1.0

- Vincoli sulla sintassi
  - preambolo XML
  - dichiarazione del DTD es:  

```
<!DOCTYPE html PUBLIC
  "-//W3C//DTD XHTML 1.0 Strict//EN"
  "xhtml1-strict.dtd">
```
  - l'elemento radice è <html>
  - il namespace del documento è  
<http://www.w3.org/1999/xhtml>
  - valgono le regole della sintassi di XML

## XHTML 1.0

- Vincoli sulla sintassi
  - il documento deve essere ben formato
  - elementi vuoti; es: <br />
  - valori degli attributi tra apici;  
es: xmlns="http://www.w3.org/1999/xhtml"
  - codifica delle entità; es: &amp;
  - numerose entità predefinite: Latin-1  
characters, Special characters, Symbols

## XHTML 1.0 e CSS

- Separazione tra struttura e presentazione
  - i tag descrivono solo la struttura
  - fogli di stile CSS: associano caratteristiche di visualizzazione agli elementi (>>)
  - foglio di stile standard del browser;  
es: testo in carattere Times 10 punti tondo

Università Roma Tre

## Elenco delle Funzionalità

- Nucleo ("Core")
  - Structure, Text, Hypertext, List
- Estensioni del testo
  - Presentation, Edit, Bidirectional Text
- Maschere
  - Basic Forms, Forms
- Tabelle
  - Basic Table, Table
- Immagini
  - Image
- Mappe
  - Client Side Map, Server Side Map
- Frame
  - Frame, Target, Iframe
- Fogli di stile
  - Stylesheet, Link
- Altri
  - Intrinsic Events, Metainformation, Scripting, Object, Base, Name, Legacy

Sistemi Distribuiti aa 2003-2004 25

Università Roma Tre

## Dichiarazione Documento XHTML 1.0

- Vincoli sulla sintassi di XHTML 1.0

```
<?xml version="1.0" ?>
<!DOCTYPE html
  PUBLIC "-//W3C//DTD XHTML Basic 1.0//EN"
    "DTD-xhtmlbasic/xhtml-basic10.dtd">
<html xmlns="http://www.w3.org/1999/xhtml">
...
</html>
```

Sistemi Distribuiti aa 2003-2004 26

Università Roma Tre

## DTD di un Documento XHTML

Sistemi Distribuiti aa 2003-2004 27

Università Roma Tre

## Struttura di un Documento XHTML

```
<?xml version="1.0" ?>
<!DOCTYPE html
  PUBLIC "-//W3C//DTD XHTML Basic 1.0//EN"
    "DTD-xhtmlbasic/xhtml-basic10.dtd">
<html xmlns="http://www.w3.org/1999/xhtml">
  <head>
    <title>Esempio</title>
  </head>
  <body>
    ...
  </body>
</html>
```

Sistemi Distribuiti aa 2003-2004 28



## Composizione di Documenti

- Elementi detti di flusso (“flow”)
  - Heading
    - titoli delle sezioni del documento
  - Block
    - blocchi identificabili del documento
    - Possono contenere altri blocchi, inline e testo
  - Inline
    - porzioni di testo, contengono solo inline o testo

## Formattazione del Testo

- **Heading**
  - `<h1>`, `<h2>`, `<h3>`,
  - `<h4>`, `<h5>`, `<h6>`
- **Block**
  - `<p>`: paragrafi
  - `<pre>`: testo preformattato
  - `<address>`: indirizzi
  - `<blockquote>`: citazioni (attributo `cite`)
  - `<div>`: raggruppamento di elem. in blocco
- **Inline**
  - `<em>`: enfaticizzato
  - `<strong>`: rafforzato
  - `<br />`: interruzione di linea
  - `<q>`: citazioni brevi (attributo `cite`)
  - `<span>`: raggruppamenti
  - ...

## Esempio

```
<html xmlns="http://www.w3.org/1999/xhtml">
  <head>
    <title>Esempio</title>
  </head>
  <body>
    <h1>Titolo principale</h1>

    <p>Questo &egrave; un documento di esempio
    relativo ad <em>XHTML Basic</em> <br />
    Sono andato a capo</p>
  </body>
</html>
```

## Collegamenti Iperestuali

- Elemento `<a>` (ancora)
  - collegamenti ad altri URI
  - attributo `href` con riferimento ad URI
  - si aggiunge agli elementi *inline*

elemento	attributo	modello di contenuto
a	href ...	(#PCDATA   <i>inline</i> )

## Esempio

```
<body>
  <h1>Titolo principale</h1>
  <p>Questo &egrave; un documento di esempio
    relativo ad <em>XHTML Basic</em> <br />
    Sono andato a capo</p>
  <a href="http://www.dia.uniroma3.it">
    <strong>Sito del DIA</strong>
  </a>
  <a href="documenti/prova.html">
    Una pagina di prova
  </a>
</body>
```

## Liste

- Liste di elementi
  - Numerate (<ol>) e non numerate (<ul>)
  - Elementi <li>

elemento	attributo	modello di contenuto
ol	...	(li)+
ul	...	(li)+
li	...	(#PCDATA   flow)*

## Esempio

```
<body> ...
  <a href="http://www.dia.uniroma3.it">
    <strong>Sito del DIA</strong>
  </a>
  <a href="documenti/prova.html">
    Una pagina di prova
  </a>
  <ul>
    <li>primo elemento di una lista</li>
    <li>secondo elemento di una lista</li>
  </ul>
</body>
```

## Immagini

- Elemento <img>
  - esprime il collegamento ad un'immagine
  - formati ordinari: gif, jpg, etc.
  - si aggiunge agli elementi *inline*
  - Attributi fondamentali: *src*, *alt*, *width*, *height*

elemento	attributo	modello di contenuto
img	src alt width, height, ...	EMPTY

## Esempio

```
<body>
  ...
  <ul>
    <li>primo elemento di una lista</li>
    <li>secondo elemento di una lista</li>
  </ul>
  
</body>
```

## Tabelle

- elemento radice: **<table>**  
attributo summary per descrivere il contenuto
- didascalia dopo la radice: **<caption>**
- una o più righe (table row): **<tr>**
- fatte di una o più celle
  - celle di intestazione (table header): **<th>**
  - celle dei valori (table data): **<td>**

## Tag per le Tabelle

elemento	attributo	modello di contenuto
table	summary ...	(caption?, tr+)
caption	...	(#PCDATA   inline)*
tr	...	(th   td)+
th	colspan, rowspan, scope (row   column) ...	(#PCDATA   flow)* meno altri elementi table
td	colspan, rowspan ...	(#PCDATA   flow)* meno altri elementi table

## Tabelle: Esempio

```
<body>
  ...
  <table>
    <caption>Lezioni</caption>
    <tr>
      <th>Data</th><th>Argomento</th><th>Tipo</th>
    </tr>
    <tr>
      <td>Lun 17 feb 03</td><td>HTTP</td><td>Lezione</td>
    </tr>
    <tr>
      <td>Mar 18 feb 03</td><td>XML</td><td>Lezione</td>
    </tr>
    <tr>
      <td>Mer 19 feb 03</td><td>XML</td><td>Esercizi</td>
    </tr>
  </table>
  ...
</body>
```

Università Roma Tre

## Tabelle: Esempio

```

<table>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>C</th>
</tr>
<tr>
<td>xxx</td>
<td>yyy</td>
<td>zzz</td>
<td>ww</td>
</tr>
<tr>
<td colspan="2">due righe</td>
<td>123</td>
<td rowspan="3">tre col.</td>
</tr>
<tr>
<td colspan="2"></td>
<td></td>
</tr>
<tr>
<td colspan="2"></td>
<td></td>
</tr>
</table>
    
```

Sistemi Distribuiti aa 2003-2004 45

## Esempio Completo

```

<?xml version="1.0" ?>
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML Basic 1.0//EN"
"DTD-xhtmlbasic/xhtml1-basic1.0.dtd">
<html xmlns="http://www.w3.org/1999/xhtml">
<head>
<title>Esempio</title>
<meta name="author" content="Pippo" />
<meta name="keywords" content="HTML" />
</head>
<body>
<h1>Titolo principale</h1>
<p>Questo &grave; un documento
di esempio relativo ad <em>HTML
Basic</em> <br />
Sono andato a capo. Una immagine
&grave; stata  immersa in questo paragrafo;
nota che l'immagine &grave; un oggetto
flottante (prova a cambiare le dimensioni
orizzontali della finestra).
</p>
Mettiamo anche un link utile: il
<a href="http://www.w3.org">
<strong>Sito del Consorzio W3C</strong>
</a>. E ora un link locale ad
<a href="prova.html">
una pagina di prova
</a>.
</html>
    
```

Sistemi Distribuiti aa 2003-2004 46

## Esempio Completo (cont)

```

<p>
Una lista con due elementi:
<ul>
<li>primo elemento di una lista</li>
<li>secondo elemento di una lista</li>
</ul>
</p>
<table>
<caption>Lezioni</caption>
<tr>
<th>Data</th><th>Argomento</th><th>Tipo</th>
</tr>
<tr>
<td>Lun 17 feb 03</td><td>HTTP</td><td>Lezione</td>
</tr>
<tr>
<td>Mar 18 feb 03</td><td>XML</td><td>Lezione</td>
</tr>
<tr>
<td>Mer 19 feb 03</td><td>XML</td><td>Eser.</td>
</tr>
</table>
</body>
</html>
    
```

Sistemi Distribuiti aa 2003-2004 47

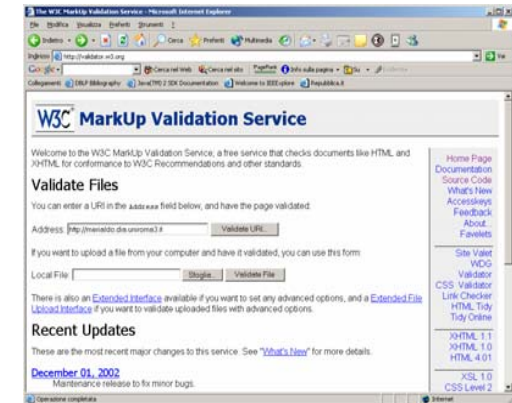
## Esempio Completo

Sistemi Distribuiti aa 2003-2004 48

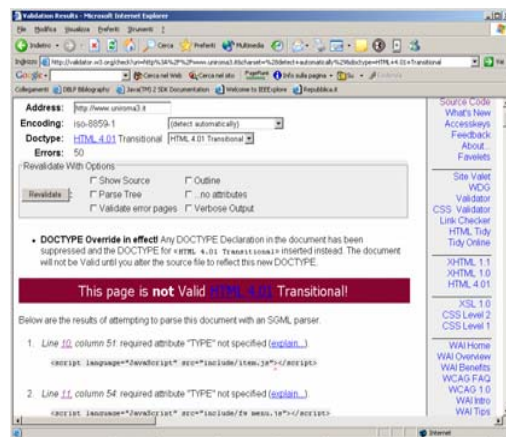
## Convalida del Codice

- Il browser non effettua convalide
- E' consigliato verificare la correttezza
  - operazione noiosa ma utile
- Validatori
  - <http://validator.w3.org>, validatore in linea
  - ...

## Validatore



## Validatore



## Termini della Licenza

- Questo lavoro:
  - si basa su materiale originale preparato dal prof. G. Mecca della Università della Basilicata nell'ambito del progetto Diogene (<http://www.difa.unibas.it/gmecca/diogene>), e successivamente adattato dall'ing. Paolo Merialdo della Università Roma Tre.
  - viene concesso in uso secondo i termini della licenza "Attribution-ShareAlike" di Creative Commons.
- Per ottenere una copia della licenza, è possibile visitare <http://creativecommons.org/licenses/by-sa/1.0/> oppure inviare una lettera all'indirizzo Creative Commons, 559 Nathan Abbott Way, Stanford, California 94305, USA.