

Corso di Informatica Generale 1 – IN1

Linguaggio HTML

Marco Liverani
(liverani@mat.uniroma3.it)

Sommario

- Breve cronologia del World Wide Web
- Introduzione e contestualizzazione
- Definizione della struttura del documento
- Link ipertestuali
- Maschere di inserimento dati (*form*)
- Formattazione tipografica
- Fogli di stile (*cascading style sheet*)
- Linee guida per la creazione di documenti per il Web

Breve cronologia del WWW

1

- **1980:** Tim Berners-Lee (CERN) sviluppa il programma “*Enquire-Within-Upon-Everything*” che consentiva di effettuare link tra diversi nodi.
- **1989:** Tim Berners-Lee diffonde due documenti per raccogliere opinioni sul suo lavoro presso il CERN.
- **1990:** Il documento di TBL viene riformulato con l'appoggio ufficiale del CERN; viene coniato il nome World Wide Web.
- **1991:** Sviluppo dei primi client ed apertura del WWW server del CERN.
- **1992:** Sviluppo del client con interfaccia X; viene rilasciata la lista dei primi 26 server WWW.
- **1993:** Viene rilasciato il primo browser per Macintosh; viene rilasciato X-Mosaic di Marc Andreessen (NCSA); i server HTTP sono circa 50.

Dicembre 2004

M. Liverani - Linguaggio HTML

3

Breve cronologia del WWW

2

- **1994:** Viene fondata la Mosaic Corporation (poi Netscape Corp.); i siti WWW sono 1.500. A dicembre si tiene il primo meeting del “W3 Consortium” presso il MIT; il CERN, a causa di problemi di budget, decide di sospendere il supporto del progetto WWW.
- **1995:** Viene fondata la “Web Society” a Graz (Austria): partecipano la Technical University di Graz, il CERN, l'Università del Minnesota e l'INRIA (Francia).

Dicembre 2004

M. Liverani - Linguaggio HTML

4

Elaborazione automatica di testi

- Elaborazione di documenti mediante il computer... Cosa significa? Qual è lo scopo?
 - **Comporre** testi con l'ausilio della macchina (è più comodo)
 - Definire l'**aspetto del testo** con qualità tipografica
 - **Riprodurlo** in più copie rapidamente
 - **Modificare il testo** o integrarlo successivamente
 - Memorizzare, **archiviare** documenti
 - Eseguire delle **ricerche** su un archivio documentale
 - **Trasmettere, condividere** documenti con altri

Dicembre 2004

M. Liverani - Linguaggio HTML

5

Strumenti – Word processor

- Il *word processor* è una delle applicazioni più diffuse sui personal computer.
- Consente di comporre e modificare un documento, consente di archiviare, stamparlo.
- Il documento può essere condiviso con chi utilizza lo **stesso programma** (e forse anche la stessa macchina, lo stesso sistema operativo, ...).
- Devo conoscere il **formato dei dati** per poter costruire strumenti di ricerca di informazioni sull'archivio documentale

Dicembre 2004

M. Liverani - Linguaggio HTML

6

Strumenti – Linguaggi di marcatura

- Il linguaggio TEX (e LATEX) permette la costruzione di documenti secondo un *formato aperto* (noto a tutti), aumentando così la possibilità di condivisione e scambio
- Necessita di un “*motore di rendering*” (un programma) per poter produrre una copia leggibile del documento
- Non permette di legare le informazioni secondo una struttura non rigidamente gerarchica e sequenziale

Dicembre 2004

M. Liverani - Linguaggio HTML

7

HyperText Markup Language

- È un “*linguaggio di marcatura*” del testo:
 - Non è un linguaggio di programmazione
 - È costituito da un insieme di parole chiave (*tag* o marcatori) con cui identifica porzioni del documento
 - È nato con l’obiettivo di consentire la definizione di *documenti ipertestuali distribuiti* su una rete di computer
- HTML viene definito come applicazione di *SGML* (*Standard Generalized Markup Language*)

Dicembre 2004

M. Liverani - Linguaggio HTML

8

Tag per il mark-up del testo

- I tag consentono di contrassegnare (marcare) una porzione del documento per associargli una determinata caratteristica (strutturale, tipografica, ecc.).
- I tag sono costituiti da un elemento di **inizio-marcatura** ed un elemento di **fine-marcatura**; i tag sono *case insensitive*.
- Gli elementi di inizio marcatura possono essere caratterizzati da **attributi** a cui è associato un particolare **valore**:

```
<tag attributo1="valore1" attributon="valoren"> ... </tag>
```

- Esempio:

```
<p align="center">Paragrafo del documento</p>
```

Dicembre 2004

M. Liverani - Linguaggio HTML

9

Entità

- Per garantire la massima **compatibilità** e **trasportabilità** del documento, HTML utilizza la codifica ASCII standard a 7 bit (caratteri da 1 a 127)
- Per rappresentare **simboli speciali** o **caratteri nazionali** utilizza alcune entità
- Esempi:

```
&agrave; = "à", &egrave; = "è", &acute; = "é",
```

```
&Egrave; = "È", &uuml; = "ü", ecc.
```

```
&amp; = "&", &quot; = " ” ", &lt; = "<", &gt; = ">", ecc.
```

Dicembre 2004

M. Liverani - Linguaggio HTML

10

Caratteristiche di HTML

- Con HTML è possibile:
 1. Definire la **struttura logica** del documento
 2. Definire dei **collegamenti ipertestuali** fra documenti (*link*)
 3. Definire l'**interfaccia utente** di una applicazione web
 4. Definire la **formattazione tipografica** del testo, fornendo indicazioni utili per il *rendering*
- L'**errore di sintassi** non è previsto nell'ambito dell'elaborazione di documento HTML: ciò che non è riconducibile ad un *tag* del linguaggio viene **ignorato** e visualizzato in quanto tale.

Dicembre 2004

M. Liverani - Linguaggio HTML

11

Primo punto di vista

**HTML come strumento
per definire la **struttura** di
un documento**

Dicembre 2004

M. Liverani - Linguaggio HTML

12

Document Object Model

- Definire la struttura logica del documento significa identificarne delle componenti riconducibili ad elementi di una **struttura gerarchica** nell'ambito di un modello denominato **DOM: Document Object Model**

Dicembre 2004

M. Liverani - Linguaggio HTML

13

Struttura del documento

- Il DOM di HTML prevede la seguente struttura gerarchica:



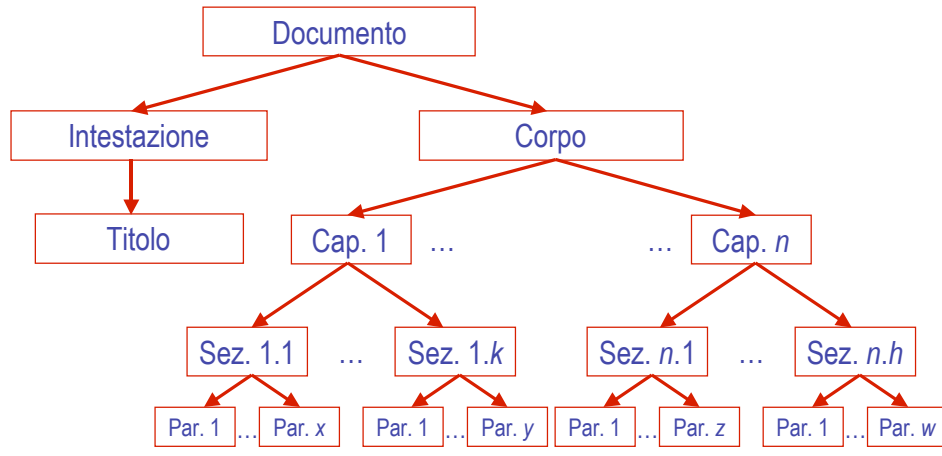
Dicembre 2004

M. Liverani - Linguaggio HTML

14

Struttura del documento

- Lo stesso modello può essere rappresentato con un albero:



Dicembre 2004

M. Liverani - Linguaggio HTML

15

Tag per la definizione della struttura



Dicembre 2004

M. Liverani - Linguaggio HTML

16

Rendering del documento

```

<html>
  <head>
    <title>Primo esempio</title>
  </head>
  <body>
    <h1>Titolo del capitolo</h1>
    <h2>Titolo della sezione</h2>
    <p>Testo del paragrafo...</p>
  </body>
</html>

```

Dicembre 2004
M. Liverani - Linguaggio HTML
17

Ancora sulla struttura “logica” del testo

- I tag che definiscono la struttura del documento consentono al programma che lo visualizza di **evidenziare graficamente** le porzioni di testo caratterizzate da un determinato “ruolo” (titoli, sottotitoli, ecc.).
- Oltre alla gerarchia dei titoli (**h1**, **h2**, ..., **h6**) esistono altri tag per caratterizzare il testo; es.: **...**: enfattizza (in qualche modo) una frase *importante* evidenziandone i caratteri.
- Anche la spaziatura tra i paragrafi e il tipo di carattere da utilizzare viene **stabilito autonomamente** dal programma che visualizza il documento.

Dicembre 2004
M. Liverani - Linguaggio HTML
18

Liste ed elenchi

- Elenchi puntati:

``

` Primo elemento`

` Secondo elemento`

``

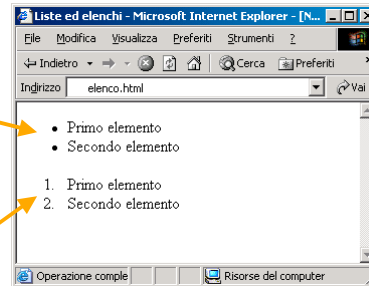
- Liste numerate:

``

` Primo elemento`

` Secondo elemento`

``



Dicembre 2004

M. Liverani - Linguaggio HTML

19

Secondo punto di vista

**HTML come strumento
per definire
documenti ipertestuali
distribuiti in rete**

Dicembre 2004

M. Liverani - Linguaggio HTML

20

Documenti ipertestuali

- I documenti HTML possono contenere riferimenti (*link ipertestuali*) ad altri documenti
- In questo modo la lettura del testo può **non** essere rigidamente sequenziale: invece di una sequenza di pagine (come in un libro) ottengo una **rete di pagine** (una sorta di *grafo* con pagine fra loro correlate)
- Il **riferimento ad una risorsa** (documento o altro) viene espresso mediante un **URL** (*Uniform resource Locator*)

Dicembre 2004

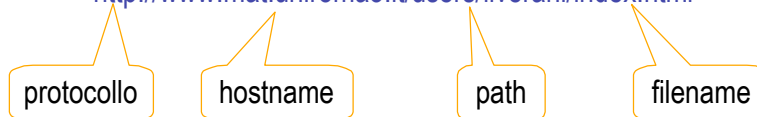
M. Liverani - Linguaggio HTML

21

Documenti distribuiti in rete

- Con un URL è possibile descrivere la **modalità** con cui deve essere raggiunto il documento correlato e la sua **collocazione**:

<http://www.mat.uniroma3.it/users/liverani/index.html>



- Il **protocollo HTTP** (*HyperText Transfer Protocol*) permette di raggiungere e trasferire documenti da una macchina (server HTTP) ad un'altra (client HTTP) su una rete TCP/IP (es.: la rete Internet)
- L'insieme di server HTTP sulla rete Internet costituisce il **World Wide Web**

Dicembre 2004

M. Liverani - Linguaggio HTML

22

Tag per la creazione di link

- Il tag per la definizione di riferimenti ipertestuali (*hypertextual reference*) è
`testo...`
- Il riferimento ipertestuale può spingersi fino ad un punto particolare del documento indirizzato da un URL: possono essere definite delle "ancore" interne al documento per indicare punti di aggancio di un link:

`testo...`

e

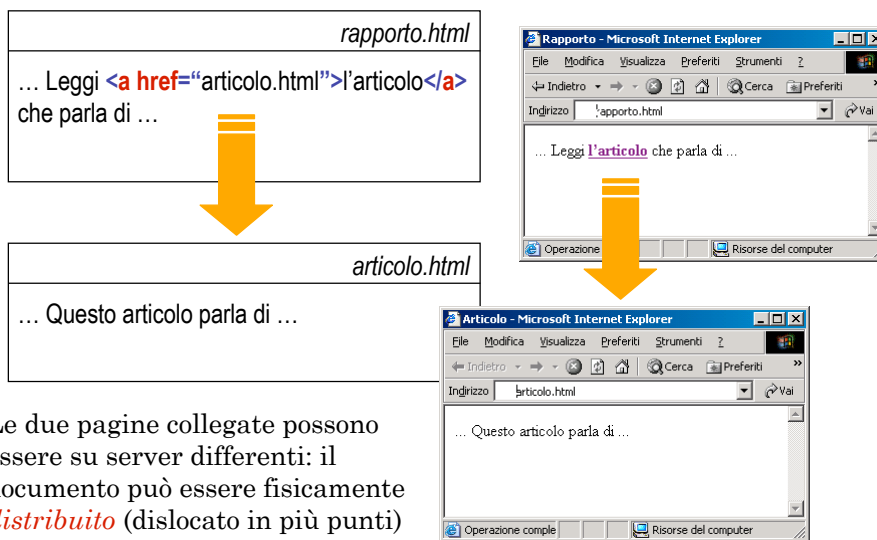
`altro testo...`

Dicembre 2004

M. Liverani - Linguaggio HTML

23

Collegamenti ipertestuali



Dicembre 2004

M. Liverani - Linguaggio HTML

24

Altri tipi di link

- Definendo un link è possibile indicare il **protocollo** con cui potrà essere raggiunta la risorsa collegata (documento, file, ...).
- Se il protocollo è differente da HTTP (es.: FTP, MAILTO, ...) verrà attivato un **programma adatto** a trattarlo (a volte lo stesso browser gestisce più protocolli).
- Se il tipo di documento richiesto non può essere visualizzato dal browser (es.: documento Word, PDF, Postscript, ecc.), tipicamente viene attivato un programma adeguato, ovvero il file viene salvato sul computer.

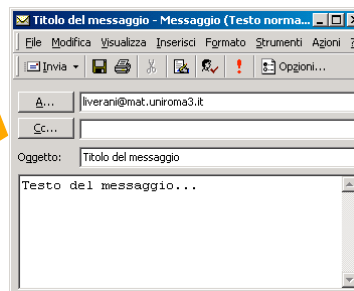
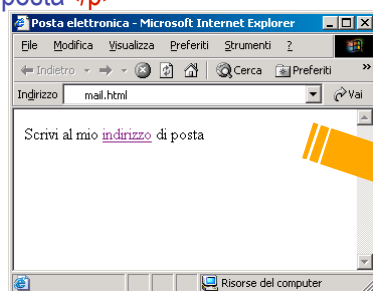
Dicembre 2004

M. Liverani - Linguaggio HTML

25

Altri tipi di link

- Esempi:
 - `<p>Scarica il file`
`file.zip...</p>`
 - `<p>Scrivi al mio`
`indirizzo di`
`posta</p>`



Dicembre 2004

M. Liverani - Linguaggio HTML

Terzo punto di vista

HTML come strumento per definire **interfacce utente** per applicazioni web

Dicembre 2004

M. Liverani - Linguaggio HTML

27

Form di inserimento dati

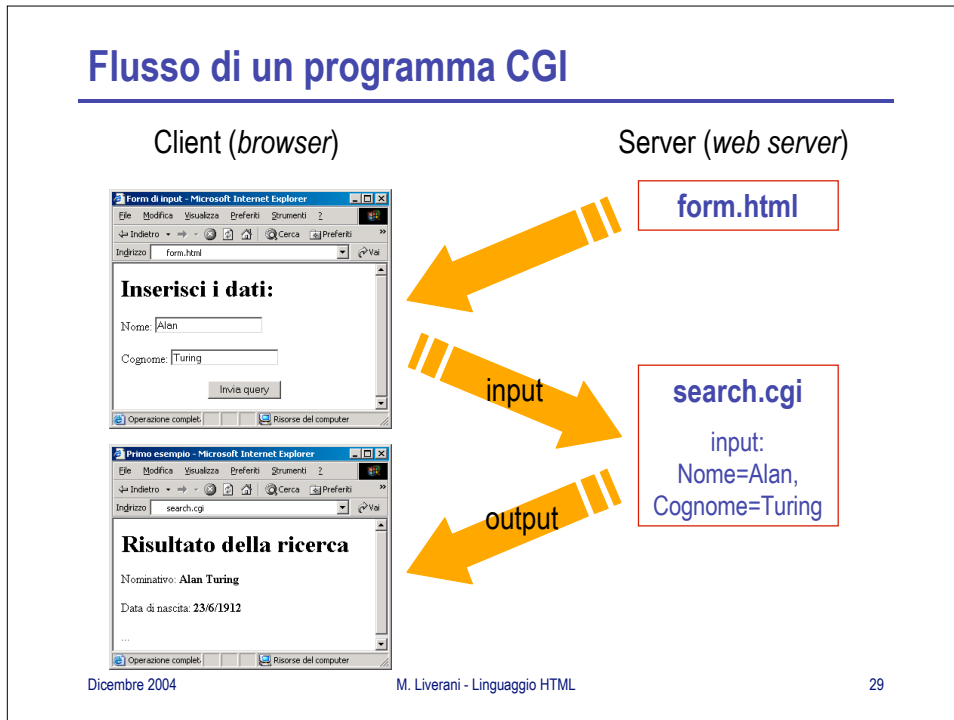
- Le pagine web possono essere usate anche come “**interfaccia**” verso un programma applicativo (es.: **programmi CGI**).
- Il programma produrrà un **output** in formato HTML per poterlo visualizzare nel browser e acquisirà **l'input** da parte dell'utente attraverso delle maschere di inserimento dati denominate **form**.
- Una form è un insieme di **campi di input** caratterizzati da un **nome** (univoco) ed un **tipo**.
- Quando l'utente ha terminato la compilazione della form viene eseguito il programma CGI passandogli come input i dati inseriti dall'utente nella form.

Dicembre 2004

M. Liverani - Linguaggio HTML

28

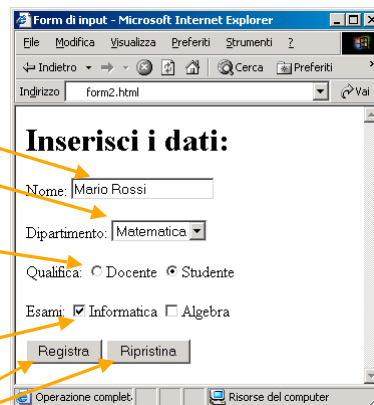
Flusso di un programma CGI



Form di inserimento dati

```

<html>
<head><title>Form di input</title></head>
<body>
<h1>Inserisci i dati:</h1>
<form action="search.cgi" method="GET">
<p>Nome: <input type="text" name="nome"></p>
<p>Dipartimento: <select name="dip">
<option value="mat">Matematica
<option value="fis">Fisica
</select></p>
<p>Qualifica: <input type="radio" name="qualifica"
value="prof" checked>Docente <input type="radio"
name="qualifica" value="stud">Studente </p>
<p>Esami: <input type="checkbox" name="esami"
value="in1">Informatica <input type="checkbox"
name="esami" value="al1">Algebra</p>
<p><input type="submit" value="Registra">
<input type="reset" value="Ripristina"></p>
</form>
</body>
</html>
    
```



Quarto punto di vista

HTML come strumento per definire l'aspetto tipografico di un documento

Dicembre 2004

M. Liverani - Linguaggio HTML

31

Formattazione tipografica

- È possibile fornire delle *indicazioni* per il *rendering* del documento HTML (per la sua rappresentazione tipografica).
- Tali indicazioni **possono essere trascurate** dal programma visualizzatore, in modo da adattare al meglio la visualizzazione della pagina allo strumento fisico su cui viene eseguito il rendering (*browser*).

Dicembre 2004

M. Liverani - Linguaggio HTML

32

Esempi di formattazione tipografica elementare

The screenshot shows a browser window titled "Primo esempio - Microsoft Internet Explorer" displaying the file "formati.html". The rendered content includes:

- grassetto (bold)** (linked from `grassetto (bold)`)
- sottolineato (underlined) (linked from `<u>sottolineato (underlined)</u>`)
- corsivo (italic)* (linked from `<i>corsivo (italic)</i>`)
- typewriter (linked from `<tt>typewriter</tt>`)
- grande (linked from `<big>grande</big>`)
- piccolo (linked from `<small>piccolo</small>`)
- x^n (apice) (linked from `xⁿ (apice)`)
- x_k (pedice) (linked from `x_k (pedice)`)

Dicembre 2004

M. Liverani - Linguaggio HTML

33

Caratteri, font, colori

- È possibile definire in dettaglio il tipo di carattere che deve essere utilizzato:
`...`
- Questo rende poco "trasportabile" il documento: è legato alla disponibilità di caratteri e colori sulla macchina dell'utente
- I colori vengono definiti specificandone le componenti **RGB** (*red, green, blue*) in notazione esadecimale (valori da 00 a FF):

$$\#rrggbb$$
- Esempio: **#AB205B**

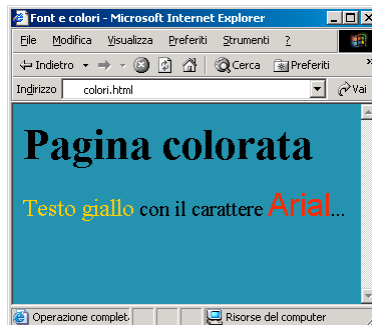
Dicembre 2004

M. Liverani - Linguaggio HTML

34

Caratteri, font, colori

```
<html>
<head>
  <title>Font e colori</title>
</head>
<body bgcolor="#3090AA">
  <h1>Pagina colorata</h1>
  <p>
    <font color="#FFCC00" size="+1">Testo giallo</font> con il carattere
    <font face="Arial" color="#FF0000" size="+2">Arial</font> ...
  </p>
</body>
</html>
```



Dicembre 2004

M. Liverani - Linguaggio HTML

35

Immagini

- È possibile includere *immagini* in un documento HTML. L'immagine verrà visualizzata compatibilmente con le capacità del browser e dell'ambiente ospite

```

```

- Esempio:

```

```



Dicembre 2004

M. Liverani - Linguaggio HTML

36

Fogli di stile

- Negli anni HTML si è arricchito di numerosi tag di caratterizzazione grafica degli elementi della pagina fino a **perdere la sua originaria eleganza e portabilità**.
- Per semplificare la costruzione di pagine HTML con una forte connotazione tipografica e cromatica, ritornando all'obiettivo originale del linguaggio (definire una struttura logica del documento) sono stati introdotti i **fogli di stile (Cascading Style Sheet – CSS)**.
- Nel CSS vengono definite le caratteristiche tipografiche dei tag presenti nella pagina.

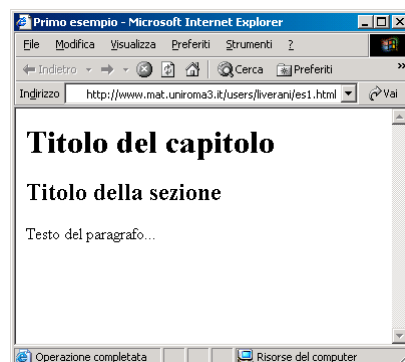
Dicembre 2004

M. Liverani - Linguaggio HTML

37

Fogli di stile

```
<html>
  <head>
    <title>Primo esempio</title>
  </head>
  <body>
    <h1>Titolo del capitolo</h1>
    <h2>Titolo della sezione</h2>
    <p>Testo del paragrafo...</p>
  </body>
</html>
```



Dicembre 2004

M. Liverani - Linguaggio HTML

38


Fogli di stile

```

H1 {
  font-family: Arial, Helvetica, sans-serif;
  text-align: center;
  font-weight: normal;
  border-color: #5588bb;
  border-style: solid;
}

p {
  font-family: Times New Roman, serif;
  text-align: justify;
  color: #990000;
  font-weight: bold;
}

body {
  background: #FFCC66;
}
        
```



```

<html>
<head>
<title>Primo esempio</title>
<link rel="stylesheet" type="text/css" href="style.css">
</head>
<body>
  <h1>Titolo del capitolo</h1>
  <h2>Titolo della sezione</h2>
  <p>Testo del paragrafo...</p>
</body>
</html>
        
```

Dicembre 2004

Conclusioni

Indicazioni per la creazione di “buoni” documenti HTML

Dicembre 2004

M. Liverani - Linguaggio HTML

40

Indicazioni per la creazione di pagine HTML 1

- Riportare sempre un **riferimento all'autore** o al responsabile delle informazioni contenute nella pagina (*webmaster*), possibilmente anche con un link ad un indirizzo di posta elettronica.
- Indicare la **data di emissione** del documento e lo stato delle informazioni in esso contenute (definitive, bozza, ecc.).
- Meglio creare un link a **risorse presenti altrove** in Rete, anziché copiarle sul nostro sito.
- **Rendere omogenee e coerenti le pagine** di uno stesso sito, anche mediante fogli di stile e icone di navigazione, per evitare che l'utente si perda dopo poche pagine.

Dicembre 2004

M. Liverani - Linguaggio HTML

41

Indicazioni per la creazione di pagine HTML 2

- Prima di iniziare lo sviluppo di un sito è bene progettare uno **schema della struttura** che assumerà il sito stesso (inizialmente un albero a cui poi aggiungere link "trasversali").
- Non creare documenti **troppo lunghi** (lenti da trasferire); meglio suddividere il documento in capitoli su file separati e collegati tra loro mediante indici ipertestuali, icone di navigazione ed *hotword*.
- **Non segmentare eccessivamente** il documento in pagine piccole: la navigazione ipertestuale è faticosa se il passaggio da una pagina ad un'altra è ingiustificato.

Dicembre 2004

M. Liverani - Linguaggio HTML

42

Indicazioni per la creazione di pagine HTML 3

- **Esplicitare sempre il contesto** in cui si colloca ogni pagina del sito: il lettore potrebbe essere giunto direttamente su una certa pagina senza aver seguito una sequenza da noi immaginata.
- Il titolo del documento deve essere breve, ma deve anche **richiamare il contesto** in cui è inserita la pagina.
- Il lay-out della pagina non deve essere progettato basandosi esclusivamente sull'output prodotto da un particolare browser: meglio effettuare un test del documento su più browser di produttori diversi.

Dicembre 2004

M. Liverani - Linguaggio HTML

43

Indicazioni per la creazione di pagine HTML 4

- Non utilizzare caratteristiche tipiche di un particolare client in modo esclusivo, senza offrire una seconda possibilità agli utenti che non potranno utilizzare o vedere determinati oggetti (font specifici, applet Java, immagini prive del *tag* ALT, *frame*, Java-script, controlli ActiveX, suoni in background o altri oggetti che richiedano particolari *plug-in*).

Dicembre 2004

M. Liverani - Linguaggio HTML

44

Riferimenti informativi

- W3 Consortium “*HTML home page*”:
<http://www.w3.org/MarkUp/>
- W3 Consortium “*Cascading Style Sheet home page*”: <http://www.w3.org/Style/CSS/>
- W3 Consortium “*Hypertext Transfer Protocol*”:
<http://www.w3.org/Protocols/>
- Tim Berners-Lee “*Style Guide for online hypertext*”:
<http://www.w3.org/Provider/Style/>
- Kevin Werbach, “*The Bare Bones Guide to HTML*”:
<http://werbach.com/barebones/> (in italiano:
http://werbach.com/barebones/it/it_barebone.html)