

2. Limbajul HTML

HTML (Hypertext Markup Language) este un set de marcate de tip text care se inserează pentru a specifica informații despre formatare și despre legături. Fișierele în format HTML au extensia implicită *.html* sau *.htm*. Toate marcatele HTML încep cu semnul < și se termină cu >. Aceste marcate sunt interpretate de către programe specializate printre care se numără browser-ele. Ca browser-e mai cunoscute putem enumera Internet Explorer și Netscape Navigator.

2.1. Noțiuni HTML de bază

Orice fișier HTML conține la început marcajul <HTML> care specifică faptul că fișierul respectiv va fi interpretat ca fișier HTML. Marcajul este de tip pereche <HTML>, </HTML>.

Trecerea la o linie nouă sau la un paragraf nou se face utilizând marcatele
 și <P>. Browser-ul nu ține cont de formatarea în linii și paragrafe pe care a avut-o textul scris cu un editor de texte (NOTEPAD)

2.1.1. Împărțirea documentului în secțiuni

Fișierele HTML sunt împărțite deseori în două secțiuni: antetul și corpul fișierului. Aceste marcate de antet și corp nu modifică afișarea fișierului HTML dar sunt necesare pentru a ușura prelucrarea ulterioară a fișierului. Marcajul de antet este <HEAD>, </HEAD>, iar cel care definește corpul fișierului este <BODY>, </BODY>. În general, în antet se specifică titlul documentelor HTML, iar corpul definește conținutul paginii HTML.

Un alt marcaj foarte des utilizat este TITLE. Utilizarea acestui marcaj permite afișarea unui titlu pe bara de sus a ferestrei afișate de browser.

2.1.2. Subtitluri

Pentru evidențierea subtitlurilor se utilizează un marcaj de forma <H_i>. <H1> </H1> este cel mai mare ca dimensiune iar <H6> </H6> este cel mai mic ca dimensiune.

Rigla orizontală

Rigla orizontală este un marcaj util pentru a organiza documentul HTML în mai multe părți. Acest marcaj produce o linie subțire orizontală. Sintaxa marcajului este <HR>.

2.2. Stiluri de text

Limbajul HTML permite evidențierea anumitor porțiuni dintr-un text într-un mod asemănător cu cel al unui editor de text. Se pot formata caractere în stilul italic, bold sau underline. Pentru asta, regiunea

care se dorește scrisă cu italic se va încadra între caracterele <I> și </I>. Pentru formatare îngroșată se utilizează și . Pentru subliniere utilizăm <U> și </U>, iar pentru text tăiat se utilizează <STRIKE> și </STRIKE>

Marcajele HTML pentru stiluri de text pot fi combinate astfel încât efectele se cumulează.

Culori, dimensiuni și seturi de caractere

Limbajul HTML are posibilitatea modificării fonturilor pentru textele afișate. Astfel, se pot modifica dimensiunea, culoarea, corpul de literă a fontului utilizat. Dimensiunile fonturilor HTML variază de la 1 la 7, 1 este cel mai mic iar 7 este cel mai mare.

Pentru modificările asupra fontului utilizat se folosește marcajul cu atributele specifice. Atributele se precizează prin construcții de forma nume=valoare. Pentru a modifica dimensiunea unei porțiuni de text se va utiliza:

```
<FONT SIZE=7>...text...</FONT>
```

Modificarea relativă a dimensiunii fontului se face cu următoarea construcție:

```
text1....<FONT SIZE=+n>...text2....</FONT>
```

În exemplul de mai sus, dimensiunea fontului va crește cu n unități pentru text2 față de text1. Creșterea sau descreșterea (-n) cu n unități nu se poate face peste limitele impuse (de la 1 la 7). In mod similar, se poate specifica o culoare pentru o porțiune de text utilizând marcajul:

```
<FONT COLOR=cod_culoare> ... </FONT>
```

Pentru specificarea altor culori, care nu sunt în tabel se utilizează notația hexazecimală.

Pentru modificarea setului de caractere se utilizează marcajul împreună cu atributul FACE. Sintaxa este:

```
<FONT FACE="Arial"> ... Text....</FONT>.
```

Dacă se doresc modificări mai complexe, se pot specifica mai multe atribute în marcajul FONT.

2.3. Liste

Listele pot fi descrise cu ajutorul limbajului HTML utilizând:

- liste de tip definiție
- liste ordonate
- liste neordonate

Pentru listele de tip definiție se vor utiliza marcaje <DL> </DL> și <DT> </DT>.

Lista va fi cuprinsă între marcajele <DL> </DL>. Pentru termenul care se definește se va utiliza marcajul <DT> iar pentru descrierea definiției se va utiliza <DD>. Listele neordonate (marcate) reprezintă tipul de listă cel mai des utilizat. Lista neordonată (marcată) este cuprinsă între elementele iar fiecare element al listei este definit prin controlul . Cele două tipuri de liste pot fi combinate.

Listele numerotate sunt utilizate pentru a afișa ordinea unor operații sau anumite succesiuni care trebuie ordonate. Spre deosebire de listele neordonate elementele listei vor fi numerotate și nu marcate cu un semn distinctiv. Marcajul pentru lista ordonată este de forma și , element care era prezent și la listele neordonate.

2.4. Referințe

Referințele externe descriu legăturile către alte resurse Web (pagini html, imagini,..). Marcajul HTML pentru referințele externe este: <A>, . Pentru a stabili o referință externă acest marcaj conține atributul HREF=valoare. Valoarea specificată poate fi vidă.

Exemplu

Adresa Web a Ministerului Educației Și Cercetării este www.edu.ro. Pentru a specifica acest lucru în cadrul unui document HTML se va utiliza notația: . Efectul este că în momentul în care cursorul ajunge deasupra adresei, forma lui se schimbă într-o “mânuță”, iar la efectuarea unui click pe adresa respectivă se va efectua saltul la pagina respectivă.

Pentru a referi un obiect care nu este document Web se va utiliza un URL adecvat.

Dacă se dorește referirea unui fișier HTML de pe calculatorul propriu, atunci se va utiliza o referință în următorul mod:

. Fișierul nume_fisier.html se va găsi în același director cu pagina HTML principală. Dacă fișierul la care se dorește accesul se găsește într-un alt arbore de directoare, fișierul se va specifica sub forma:

Referințele interne se utilizează de obicei pentru a alcătui cuprinsul la documente Web de dimensiuni mari. O referință internă se specifică utilizând marcajul și .

Prin NAME="nume_referință" se specifică locația destinație a saltului. Sursa saltului se specifică prin: . Se observă că nume_referință este precedat de simbolul #.

Adăugarea imaginilor la paginile Web

Formatele permise de HTML 4 sunt GIF (**G**raphics **I**nterchange **F**ormat), JPEG (**J**oint **P**hotographic **E**xpert **G**roup). Există de asemenea și un format mai recent-formatul PNG (**P**ortable **N**etwork **G**raphic) care oferă multe din proprietățile ambelor formate. GIF permite formate pe maxim 256 culori, iar JPEG permite milioane de culori distincte într-o imagine.

Includerea imaginilor în paginile Web se face utilizând marcajul . Marcajul are ca atribut principal SRC=nume_fisier_imagine. Prin atributul SRC se precizează numele și locația imaginii. O imagine se specifică în felul următor:

```
<IMG SRC=grafic.gif>
```

Pentru browser-ele care nu pot încărca imagini, se poate preciza atributul ALT=text_alternativ, ceea ce reprezintă un text care va fi afișat în locul imaginii.

```
<IMG SRC = "image.gif" ALT= "[imaginea experimentului]"
```

Opțiuni de aliniere a imaginii

Imaginile pot fi aliniate în plan vertical față de text astfel:

ALIGN=TOP, ALIGN=MIDDLE, ALLIGN=BOTTOM

În plan orizontal imaginile pot fi aliniate stânga, dreapta, utilizând atributele ALIGN=right sau ALIGN=left. De asemenea se pot specifica dimensiunile imaginilor prin WIDTH (lățime) și HEIGHT (înălțime). Se mai pot utiliza VSPACE și HSPACE care specifică numărul de pixeli între text și imagine.

2.5. Tabele și cadre

Realizarea tabelelor în HTML este o operațiune destul de dificilă. Pentru a descrie un tabel se vor utiliza marcaje pentru fiecare celulă din tabel. Marcajele utilizate sunt:

<TABLE> </TABLE> semnalizează începutul și sfârșitul tabelului.

<TR> </TR> descrie un rând de tabel. El se poziționează la începutul și sfârșitul descrierii unui rând.

<TD> </TD> marchează începutul și sfârșitul descrierii unei celule.

Fiecare linie poate avea o aliniere proprie ca parte a marcajului <TR>. <TR ALIGN=center>. De asemenea, la marcajul <TABLE> se pot adăuga atributele:

BORDER = n – specifică lățimea zonei umbrite din jurul tabelului

CELLSPACING = n – spațierea dintre celule în pixeli

CELLPADDING=n – spațierea dintre chenar și conținutul celulei (pixeli)

WIDTH = n –lățimea dorită în pixeli sau procente

Cadre

Cadrelor au fost introduse pentru prima oară de Netscape anterior versiunii HTML 3.2. Un cadru reprezintă o porțiune dintr-o pagină Web care se comportă ca o pagină Web separată. Cadrele se specifică utilizând marcajele FRAMESET (acesta indică volumul de spațiu alocat fiecărui panou) și

FRAME (acesta indică cadrele individuale). Cu alte cuvinte, marcajul FRAME va indica pagina Web care va fi afișată în acel panou.

Marcajul FRAMESET utilizează atributele COL=x (reprezintă numărul și dimensiunea relativă a cadrelor pe orizontală) și ROW=x (reprezintă numărul și dimensiunea relativă a cadrelor pe verticală).

2.6. Formulare. Introducerea datelor

Formularele permit construirea unei pagini Web care dă posibilitatea introducerii datelor de către utilizatori. Paginile Web care conțin formulare preiau datele de la utilizator și le transmit la distanță unui server Web. Există două metode prin care informațiile pot fi transmise de la browser la serverul Web:

Metoda GET

GET este cea mai simplă și frecvent utilizată metodă de interogare. Este folosită în mod uzual pentru a accesa resurse statice precum documente HTML sau imagini. GET poate fi folosită și pentru a accesa informație în mod dinamic prin includerea parametrilor de interogare în cererea URL. De exemplu, se poate trimite un parametru "nume_cont", cu valoarea "Ionescu.Vasile", atașate la un URL ca *http://adresa?nume_cont=Ionescu.Vasile*.

Metoda POST

Metoda POST de interogare este utilizată în general pentru accesarea resurselor dinamice. Cererile POST sunt utilizate pentru transmiterea informației care depinde în mod dinamic de cerere sau atunci când o cantitate foarte mare de informații trebuie trimisă la server.

Cu metoda GET cererea este trimisă ca un șir de interogare atașat la URL. În cazul metodei POST, parametrii sunt trimisi în corpul cererii. Unele servere impun anumite restricții asupra lungimii cererii URL ceea ce duce la utilizarea metodei POST pentru interogări.

Formularele HTML sunt delimitate de marcajul <FORM> care se specifică prin <FORM ACTION=url METHOD=metoda> și de marcajul pereche </FORM>. URL-ul specifică fișierul sau aplicația de pe serverul Web de la distanță care va prelucra informațiile, iar metoda este fie GET, fie POST. De obicei, câmpurile destinate introducerii datelor se specifică prin intermediul marcajului INPUT și a unor atribute:

TYPE=tipul câmpului de intrare

Tipul câmpului de intrare poate fi:

- *text*=câmp de tip text
- *password*=câmp de text iar datele introduse de utilizator sunt afișate prin caracterul *
- *checkbox* = casetă de validare

- *hidden* = ascuns : se transmit informații programului de la distanță, iar aceste informații nu vor fi afișate pe ecranul monitorului.
- *radio* = afișează un buton cu interblocare

NAME = specifică numele simbolic al valorii câmpului

VALUE = specifică conținutul prestabilit al câmpului de tip text

SIZE = numărul de caractere al casetei de text afișate

Cele mai importante tipuri <INPUT> sunt:

SUBMIT (transmite) care determină apariția în formular a unui buton care, dacă este apăsat, transmite conținutul întregului formular la serverul de la distanță.

IMAGE – identic cu **SUBMIT**, numai că în locul butonului se specifică o imagine grafică

RESET – permite utilizatorilor să ștergă conținutul tuturor câmpurilor formularului

În exemplul de mai sus se observă că aplicația de la distanță care va prelucra datele este *formular.cgi* și se găsește la adresa <http://www.siteweb.ro/aplicatii>. Metoda de preluare a datelor este GET. Câmpurile de intrare se specifică prin INPUT. Butonul de trimitere a formularului la distanță este "Trimite" iar cel de ștergere este "Sterge textul introdus". Aceste butoane au fost specificate în marcajul HTML :

```
<INPUT TYPE=SUBMIT VALUE="Trimite">
<INPUT TYPE=RESET VALUE="Sterge textul introdus">
```

Interfața CGI

CGI (**C**ommon **G**ateway **I**nterface – Interfața comună pentru porți de acces) este un standard care specifică formatele de date utilizate pentru schimburi de informații între servere WEB, browsere și programe. Un scenariu CGI este un program scris în orice limbaj de programare, care prelucrează datele primite de la utilizator și furnizează un răspuns prin construirea unui formular HTML.

Etapele pe care le parcurge un scenariu CGI sunt:

- browser-ul solicită documentul de la serverul WEB;
- dacă documentul există, atunci acesta este furnizat browser-ului;
- utilizatorul introduce date;
- datele sunt preluate de browser și trimise la server;
- programul CGI preia datele recepționate, le prelucrează și construiește un document HTML pe care-l trimite la browser.