

HTML aneb jak psát webové stránky

1. Historie HTML a souvislost s Internetem

Historie

HyperText Markup Language (HTML) je doslova přeloženo **značkovací** jazyk pro hypertext.

V roce 1980, fyzik CERNU Tim Berners-Lee navrhl systém pro sdílení dokumentů, textů – klíčovou vlastností byla provázanost odkazy, která vedla na první specifikaci jazyku HTML v roce 1989. V roce 1990 vytvořil Berners-Lee první specifikaci a současně první webový prohlížeč, který dokázal hypertextové dokumenty zobrazovat.

Role HTML v internetu

HTML stránky tvoří od masového rozšíření Internetu do současnosti základ “pavučiny” webových stránek (WWW, www. alias World Wide Web). Jeho typické použití jsou samostatné stránky, od jednoduchých vizitek po obrovské portály, encyklopedie. Často se také používá v emailech.

Celá architektura internetových stránek je tvořena:

- Serverem; typicky počítač, který běží 24 hodin denně, 7 dní v týdnu na rychlém připojení, stará se o uložení stránek a jejich zobrazování návštěvníkům,
- komunikačním protokolem – HyperText Transfer Protocol (HTTP) (<http://www.google.com/>), jednoduchou specifikací, jak přenášet obsah po síti,
- webovým prohlížečem; softwarem, který zobrazuje přenesená data uživateli, v současné době jsou populární prohlížeče Mozilla Firefox, Internet Explorer, Opera, Chrome.

Současnost

O jazyk HTML se stará World Wide Web Consortium (W3C), společnost, která aktivně vyvíjí HTML a vydává formálně přesné specifikace (ISO/IEC 15445, W3C HTML 4.01, W3C HTML 5) a Web Hypertext Application Technology Working Group (WHATWG), sdružení lidí, kteří se o vývoj HTML a souvisejících technologií zajímají.

Současná verze HTML je 4.1 z let 1997-2000, novinkou je HTML 5 (ve formě draftu od roku 2008, ale přesto podporován moderními prohlížeči). V současném internetu se dá narazit i na stránky v HTML 3.2 z roku 1997.

Paralelně k HTML je “formálněji striktnější” jazyk XHTML (verze 1.0 z roku 2000), který vychází z HTML 4.01 a XML 1.0 a představuje paralelu k HTML 5 (XHTML 5, draft spolu s HTML 5).

2. První HTML stránka

Co budeme nutně potřebovat:

- Textový editor (notepad, PSpad, ...),
- webový prohlížeč (Firefox, Chrome, ...).

Co potřeba pro začátek nebude:

- Server,
- internetové připojení,
- WYSIWYG (What You See Is What You Get) editor.

První stránka

- 1) Vytvoříme složku HTML.
- 2) Vytvoříme podsložku pokus.
- 3) Vytvoříme soubor `test.html`.
- 4) Otevřeme soubor `test.html` a vložíme do něj následující textový obsah:

```
<html>
<head>
<title>Hello HTML</title>
</head>
<body>
<p>Hello World!</p>
</body>
</html>
```

- 5) Soubor `test.html` uložíme a otevřeme ve webovém prohlížeči.
- 6) Hurá. :)
- 7) Znovu otevřeme soubor `test.html` a upravíme ho přidáním dvou slov:

```
<html>
<head>
<title>Hello HTML</title>
</head>
<body>
<p><b>Hello</b> World!</p>
</body>
</html>
```

- 8) Soubor znovu uložíme a znovu otevřeme.
- 9) Hotovo. :)

Tedy:

- hlavním rysem HTML je tedy značkování textu (... Markup Language) ve stylu “tady chci mít tučný text”,
- stránky se dají snadno a rychle tvořit metodou pokus/omyl.

3. První HTML stránka formálněji

Tagy

HTML používá 3 hlavní typy značek (slagově “tagů”), podle Wikipedie:

- Strukturální značky, rozvrhují strukturu dokumentu, příkladem jsou odstavce (<p>), nadpisy (<h1>, <h2>). Dodávají dokumentu formu.
- Popisné (sémantické) značky, popisují povahu obsahu elementu, příkladem je nadpis (<title>) nebo adresa (<address>). Současný trend je orientován právě na sémantické značky, které usnadňují automatizované zpracovávání dokumentů a vyhledávání informací v záplavě dokumentů na webu. Vyvrcholením této snahy je v současné době jazyk XML.
- Stylistické značky, určují vzhled elementu při zobrazení, typickým příkladem je značka pro tučné písmo (). Od tohoto druhu značek se postupně upouští, trendem je používání kaskádových stylů, které vzhled popisují odděleně od obsahu dokumentu. Mezi důvody pro neuvádění těchto značek patří především to, že tyto značky jsou orientovány na prohlížení na obrazovce počítače, příliš se však nepočítá s používáním dokumentu jiným způsobem – alternativní prohlížeče pro postižené (čtečky pro slepce), v mobilních zařízeních a podobně. Kaskádové styly umožňují definovat rozdílné zobrazení pro různá zařízení.

Celé umění tvorby HTML stránek je skryté ve znalosti a použití značek. Dobrý trik je najít si nějakou stránku (např. <http://www.jakpsatweb.cz/>), kde jsou značky a jejich použití stručně a jasně popsány.

První příklad ještě jednou

Každá HTML stránka zpravidla obsahuje:

Hlavičku

```
<html>
  <head>
    <title>Hello HTML</title>
  </head>
```

Hlavička typicky určuje povahu obsahu – nadpis, jazyk, klíčová slova, grafický styl apod.

Tělo

```
<body>
  <p>
    Hello World!
  </p>
</body>
</html>
```

Tělo je hlavní část stránky – nese obsah.

Použité tagy:

- `<html></html>`

Označuje začátek a konec HTML dokumentu. Všimněme si, že je **párový**, má svůj začátek a konec, “ovlivňuje text uvnitř”. Pokud je nějaký tag párový, musíme ho zpravidla vždy uzavřít.

V HTML je většina tagů párová.

- `<head>`

Párový tag, který vymezuje hlavičku HTML dokumentu.

- `<title>`

Párový, vymezuje název dokumentu.

- `<body>`

Párový, označuje tělo dokumentu.

- `<p></p>`

Párový, označuje odstavec textu.

4. Základní tagy

Členění textu

HTML dvě a více mezer nahrazuje za mezeru jednu. To znamená, že text bez dalšího členění se “slije” v jeden dlouhý blok:

```
<html>
<head>
<title>Lorem ipsum</title>
</head>
<body>
Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor
incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud
exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute
irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla
pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia
deserunt mollit anim id est laborum.

Suspendisse potenti. Fusce venenatis facilisis nibh, quis mollis orci sagittis id.
Suspendisse a tellus leo, eu tincidunt dolor. Mauris ut nisi ante, sit amet
faucibus nibh. Aenean quis velit quis dolor fringilla faucibus vitae a urna. Duis
arcu massa, ultricies vitae venenatis et, euismod sed metus. Integer eu orci nulla.
Nullam orci diam, vestibulum quis aliquet ut, elementum vel metus.
</body>
</html>
```

Proto používáme odstavce:

```
<html>
<head>
<title>Lorem ipsum</title>
</head>
<body>
<p>Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor
incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud
exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute
irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla
pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia
deserunt mollit anim id est laborum.</p>

<p>
Suspendisse potenti. Fusce venenatis facilisis nibh, quis mollis orci sagittis id.
Suspendisse a tellus leo, eu tincidunt dolor. Mauris ut nisi ante, sit amet
faucibus nibh. Aenean quis velit quis dolor fringilla faucibus vitae a urna. Duis
arcu massa, ultricies vitae venenatis et, euismod sed metus. Integer eu orci nulla.
Nullam orci diam, vestibulum quis aliquet ut, elementum vel metus.
</p>
</body>
</html>
```

Nezáleží na tom, jestli je tag **<p>** těsně vedle textu, nebo o řádek před ním – HTML mezery ignoruje. Důležité je ale používat uzavírací značky (**</p>**) a psát tagy bez mezer (**<p>**). Používání velkých písmen uvnitř tagů není zakázáno (**<BODY>**), ale není už obvyklé.

Obecně, všechny odstavce v HTML označujeme tagy pro odstavce.

Další uživatelné značky pro formátování jsou nepárový (!) konec řádku (který se tak dá neelegantně použít místo odstavců):

```
<br>
```

Například:

```
Suspendisse potenti. Fusce venenatis facilisis nibh, quis mollis orci sagittis id.
Suspendisse a tellus leo, eu tincidunt dolor. Mauris ut nisi ante, sit amet
faucibus nibh. Aenean quis velit quis dolor fringilla faucibus vitae a urna. Duis
arcu massa, ultricies vitae venenatis et, euismod sed metus. Integer eu orci nulla.
Nullam orci diam, vestibulum quis aliquet ut, elementum vel metus.<br>
<br>
Suspendisse potenti.<br><br>
```

a opět nepárová horizontální čára:

```
<hr>
```

Odkazy

Odkazy jsou základem Hypertextu. V HTML se zapisují pomocí párové značky **<a>**, která má navíc atribut **href**, který označuje cíl odkazu. Cíl může být relativní (cesta k souboru ve stejném adresáři) nebo absolutní (kompletní adresa).

Například:

```
<a href="flip.html">flip</a>
```

```
<a href="flop.html">flop</a>
```

```
<a href="http://www.google.com">Google</a>
```

Užitečné, i když ne příliš moderní, je otevírání odkazu v novém okně pomocí atributu **target**:

```
<a href="http://www.google.com" target="_blank">Google</a>
```

Nadpisy a členění textu

Nepárové tagy, slouží pro zobrazování nadpisů různých úrovní:

- **<h1>** 1. úroveň,
- **<h2>** 2. úroveň,
- **<h3>** 3. úroveň,
- **<h4>** 4. úroveň,
- **<h5>** 5. úroveň,
- **<h6>** 6. úroveň.

Párový tagy pro centrování a text bez formátování:

- **<center>** vycentrování obsahu,
- **<pre>** text bez formátování (vhodné pro výstup z programů apod.).

Úpravy písma

Tagy pro určení formátu textu (lze kombinovat, všechny párové):

- **** tučné písmo,
- **<i>** kurzíva,
- **<tt>** neproporcionální písmo.

Obrázky

Obrázek se uvádí nepárovým tagem ****, který má několik atributů:

- **src**, relativní nebo absolutní cesta k souboru s obrázkem (podobně jako u odkazu),
- **alt**, **title**, alternativní popis; titulek,

- width, šířka obrázku (optimálně stejná, jako je velikost obrázku),
- height, výška,
- border, šířka rámečku,
- align, zarovnání obrázku (left, right).

Například:

```

```

Entity

Entity jsou speciální znaky, které nejde jinak snadno zapsat (pevná mezera – více mezer najednou, ostré závory), nejčastější:

- pevná mezera ; ,
- < < ; ,
- > > ; ,
- © © ; .

Hlavička, použití češtiny

Každý HTML 4.01 by měl obsahovat specifikaci typu dokumentu, obvykle například:

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN">
```

Tuto specifikaci je dobré uvádět v každém dokumentu jako první řádek ještě před značkou <html>, označení typu dokumentu je nepárové.

Podporu českých znaků (a všech dalších) je možné doplnit určením kódování rozšířených znaků, pro české stránky připadá v úvahu windows-1250, utf-8 a iso8859-2.

Například:

```
<meta http-equiv="content-type" content="text/html; charset=windows-1250">
```

Náš první příklad by tedy měl správně vypadat po doplnění češtiny:

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN">
<html>
<head>
<title>Hello HTML</title>
<meta http-equiv="content-type" content="text/html; charset=windows-1250">
</head>
<body>
<p>Ahoj světe!</p>
</body>
</html>
```

Seznamy

Seznamy se tvoří snadno párovou definicí typu seznamu a pak výčtem prvků v párovém tagu ``, kde typ seznamu může být

- ``, nečíslovaný seznam,
- ``, číslovaný seznam.

Tabulky

Tabulky se v HTML používají k zobrazení tabulek, ale současně také k řízení toku textu – tzv. tabulkovému designu (ve stylu např. “horní pruh menu, spodní pruh pokračování a text).

Tabulka je uvedena párovým tagem `<table>` který má atributy:

– `border`, šířka rámečku 0-1,2,3...

Uvnitř tabulky jsou jednotlivá pole, které jsou odděleny párovými tagy `<tr>` pro určení řádků a `<td>` pro určení sloupců. Sloupce se dají slučovat (`colspan`), stejně jako řádky (`rowspan`). Například:

```
<table border="1">
<tr><td><b>Leden</b></td><td>10</td><td>20</td></tr>
<tr><td><b>Únor</b></td><td>10</td><td>20</td></tr>
<tr><td><b>Březen</b></td><td>30</td><td>15</td></tr>
</table>
```

```
<table border="1">
<tr><td>A</td><td>B</td><td rowspan="2">C</td></tr>
<tr><td colspan="2">D</td></tr>
<tr><td colspan="3">E</td></tr>
</table>
```

5. Kaskádové styly

Jiný přístup k formátování obsahu, forma (CCS) je oddělená od obsahu (HTML).

6. Validátory

- <http://validator.w3.org/>

7. Nekompatibilita webových prohlížečů

Každý z výrobců implementuje standardy po svém ...

8. Dynamické jazyky a databáze

- redakční systémy
- eshopy
- blogy

- PHP, ASP .NET
- SQL

9. Další technologie

- flash
- javascript

Zdroje

- <http://en.wikipedia.org/wiki/HTML>
- http://en.wikipedia.org/wiki/World_Wide_Web_Consortium
- <http://www.jakpsatweb.cz/>
- http://en.wikipedia.org/wiki/Lorem_ipsum

2010 Ondřej Čečák <ondrej@cecak.cz>, <http://www.cecak.cz/dyzajn/>