

# HTML

## Počátky HTML

## HTML a jeho nástroje

## Struktura dokumentu

## Další vývoj značkovacích jazyků

## Základní charakteristika XHTML

### Počátky HTML

V této kapitole nejde ani tak o historii, ale o **pochopení HTML na základě historického vývoje**.

- První definici **jazyka HTML** vytvořil v roce **1991 Tim Berners-Lee** jako součást projektu WWW v Evropském centru jaderného výzkumu (**CERN**) v Ženevě. HTML v jeho pojetí umožňoval rozčlenit text do několika logických úrovní, použít několik úrovní zvýraznění textu a opatřit text odkazy a obrázky. Jeho WWW-software zahrnoval prohlížeč i integrovaný editor WWW-stránek. T.B.L. nepředpokládal, že každý autor WWW bude muset znát jazyk HTML.
- Američtí autoři prvního prohlížeče WWW-stránek **Mosaic** považovali za příliš obtížné implementovat do prohlížeče i editor. Důsledkem toho je, že oproti představě T.B.L se stala znalost HTML pro autory nutností.
- T.B.L. pokračoval ve vývoji HTML a zavedl další rozšíření – zejména formuláře a tabulky. V polovině 90. let minulého století se na vývoji HTML podílejí významně tvůrci prohlížečů. V roce 1996 se do procesu koordinace vývoje HTML zapojuje **konsorcium W3C**, jehož členem jsou mj. významné softwarové firmy. Výsledkem jeho snahy je **standard HTML 3.2**.
- Vedle konsorcia **World Wide Web Consortium (W3C)** ([www.w3.org](http://www.w3.org)) se vývoji internetu věnuje i pracovní skupina Internet Engineering Task Force (IETF) ([www.ietf.org](http://www.ietf.org)), která má záběr ještě širší než W3C a je odpovědná za definici a správu veškerých aspektů internetových technologií.
- Výsledkem doporučení W3C jsou v roce 1996 i **kaskádové styly (CSS)** a v roce 1997 i návrh **HTML 4.0**.

### HTML a jeho nástroje

**Jazyk HTML** (*HyperText Markup Language*) je jazyk pro specifikaci rozvržení dokumentu a hypertextových odkazů. Definuje syntaxi a rozmístění speciálních vložených prvků, které se v dokumentu přímo nezobrazují, ale které řídí způsob zobrazení obsahu dokumentu (textu, obrázků a ostatních podpůrných médií).

Posílat „hotové dokumenty“ typu např. Wordu po internetu jakožto www stránky by bylo obtížné, neboť dokument s kompletním formátem by byl rozsáhlý a dlouho by se natahoval, navíc uživatelé internetu mají zcela rozdílné technické a programové prostředky – různou grafiku monitoru, odlišné operační systémy (Windows 9x či 200x, Unix, Linux, OS/2, Macintosh) apod., takže zaručit, aby dokument se zobrazil přijatelným způsobem ve všech prostředích, by bylo velmi složité.

Základem www stránky je proto „holý text“ doplněný formátovacími značkami v HTML. Prohlížeč www stránek formátovací značky interpretuje a vytváří konečnou podobu dokumentu, která se může lišit v závislosti na prohlížeči, grafice počítače, operačním systému apod. Pokud je formátovací značka chybná či pokud ji prohlížeč nezná, prostě ji ignoruje. Zdá se to

nedokonalé, avšak je to rozhodně lepší než odmítnutí zobrazení stránky či ohlášení chyby ve snaze zobrazit stránku naprosto přesně.

Soubor s www stránkou má **příponu .html** nebo **.htm**.

Základem HTML jsou značky zvané **tagy** (štítky). Jsou uzavřeny do znaků <>, znak < představuje začátek tagu a znak > jeho konec. Koncový tag má po znaku < znak /.

**Tagy lze rozdělit na:**

- **Párové** – přičemž první tag je otevírací a aktivuje určitý formát až do doby, kdy je nalezen ukončovací tag.
- **Nepárové** – nemají uzavírací tag; např. pro vložení vodorovné čáry, obrázku apod.

Tagy mohou obsahovat **atributy**, kterým se jazyce HTML přiřazují jednotlivé **hodnoty**. Hodnoty jsou od atributů odděleny znakem rovná se (=), jednoslovné hodnoty nebo čísla (bez mezer) se zapisují bez uvozovek či apostrofů, víceslovné hodnoty se zapisují uzavřené do uvozovek nebo apostrofů. Atributy se oddělují „bílymi znaky“ (tabulátory, mezery, oddělovače řádků) a jejich pořadí může být zcela libovolné.

**Komentáře** prohlížeč nezobrazuje. Zapisují se mezi dvojici značek <!-- a -->.

**Tvůrce HTML potřebuje:**

- **HTML editor** – pro vytváření stránek
- **prohlížeč www-stránek** – pro prohlížení výsledků práce – doporučuje se užít pro testování více prohlížečů
- **připojení na internet** (není vždy bezpodmínečně nutné, ale je potřebné pro publikaci prací na internetu a pro získávání různých zdrojů a ověřování použitých odkazů).

### HTML editory

V zásadě lze užít:

- **Editor neformátovaného textu** (např. **Poznámkový blok**) – nutno psát úplně všechny tagy, pro začátečníka to není vhodný způsob, profesionálové často užívají pro rychlou úpravu stránek.
- **HTML editor na úrovni tagů** – **není WYSIWYG**, kontroluje syntaxi, napomáhá při psaní zdrojového textu. Výhodou je možnost velmi přehledného zdroje, barevné odlišení tagů od běžného textu, uživatel stále musí znát HTML, avšak již ne tak důkladně jako u editoru neformátovaného textu. Typickým představitelem je **HomeSite**.
- **HTML editor na úrovni výsledného dokumentu** – **WYSIWYG editor**, který je již blízký editoru typu Word, uživatel vkládá prvky stránky přičemž může sledovat neustále zdrojový text v HTML a některé úpravy provádět přímo v HTML. Tyto editory slouží též správě internetového serveru, takže není třeba užívat FTP komunikace se serverem. Typickým příkladem je **Microsoft FrontPage**. Používá se často v profesionální oblasti tvorby www stránek.
- **Export do HTML přímo z textového editoru typu Word** – uživatel nemusí znát HTML vůbec, výsledný zdrojový kód však může být rozsáhlý, neefektivní a nepřehledný. Exportem se (přes stále zdokonalování) někdy nepodaří převést úplně všechny prvky dokumentu a musí přijít částečná oprava pomocí zdrojového kódu, takže určité znalosti HTML se uživatel ani zde zpravidla zcela nevyhne.

## Struktura dokumentu

Základní značky vymežující oblasti HTML souboru:

Tag	Význam	Párový	Výskyt
<b>html</b>	začátek HTML dokumentu	ano	na začátku souboru
<b>head</b>	hlavička stránky	ano	na začátku souboru
<b>body</b>	tělo stránky + definice pozadí	ano	za <head>
<b>&lt;!-- --&gt;</b>	poznámka	ano	kdekoliv
<b>!doctype</b>	specifikace DTD	ne	úplně na začátku souboru

O specifikaci DTD bude zmínka v závěru. Popis atributů body zde není uveden (viz referenční příručka jazyka).

V následující části bude uveden přehled prvků jazyka HTML s uvedením textu, který příslušný prvek popisuje.

- ▶ **Práce s textem** – text *Formátování textu v (X)HTML*
- ▶ **Oddělovače, obrázky a multimédia** – text *Grafika na www*
- ▶ **Odkazy** – text *Odkazy v HTML a CSS*
- ▶ **Formátované seznamy** - text *Formátování textu v (X)HTML*
- ▶ **Formuláře** – text *PHP a formuláře*
- ▶ **Tabulky** – text *Tabulky v HTML a CSS*

**Rámy** nejsou v textech zahrnuty. Případného zájemce odkazují na referenční příručku jazyka.

## Další vývoj značkovacích jazyků

- 1998 - Objevuje se jazyk **XML a DHTML** (viz text *Internetové technologie*)
- 1999 - Vývoj **HTML** se završuje konečnou verzí **4.01**
- 2000 - Spojením **XML 1.0 a HTML 4.01** vzniká rozšířená verze HTML – jazyk **XHTML 1.0**. Vývoj tabulek kaskádových stylů pokračuje verzí **CSS Level 2** a vzniká modifikace značkovacího jazyka pro tvorbu stránek WAP – jazyk **WML 1.0**

## Základní charakteristika XHTML

Jazyk XHTML vznikl **spojením HTML 4.01 a XML 1.0** a upravených typových definic dokumentu (**DTD**). Jde v podstatě o „přidání inteligence“ do jazyka HTML a zpřísnění syntaktických pravidel. Je možno definovat i vlastní elementy podle pravidel jazyka XML.

- Dokument začíná obdobně jako dokument HTML např.:

```
<?xml version="1.0"? encoding="windows-1250"?>
```

- **DTD** - Vyžaduje se buď striktní definice (**strict**), což znamená, že z možných způsobů jak něco napsat používat důsledně způsob nejnovější, tj. dle HTML 4.01. Vzhled stránky lze upravovat pouze pomocí CSS. Existuje však také možnost použít přechodovou definici, která se přibližuje současnému standardu (**transitional**) Lze používat vše z HTML 4,0, nikoli však rámy. Poslední verzí je **DTD frameset** – což je přechodová verze DTD s rámy.
- Jazyk XHTML je *case sensitiv*, rozlišuje tedy malá a velká písmena. Všechny elementy, atributy a parametry je nutno zapisovat malými písmeny. Jediné co se píše velkým písmenem je <!DOCTYPE> - tedy definice DTD (viz předchozí odstavec).
- Všechny elementy musí mít ukončovací tagy nebo je nutno psát nepárové tagy speciálním způsobem např. místo <hr> se píše <hr/>.
- Elementy je nutno správně vnořovat, což sice vyžaduje už jazyk HTML, avšak většina prohlížečů si poradí, i když to není důsledně dodržováno.
- Hodnoty atributů musí být důsledně v závorkách.
- Všechny atributy musí mít explicitní hodnoty, tedy i atributy, které normálně hodnotu nemají, protože představují stav zapnuto/vypnuto, takže je nutno psát např. `noshade="noshade"`.

### Postup převodu stávajících stránek z HTML na XHTML

- Všechny názvy tagů a atributů převést na malá písmena
- Hodnoty atributů uzavřít do uvozovek
- Ukončit všechny nepárové tagy
- Zajistit nekřížení tagů
- U každého atributu napsat jeho hodnotu
- Vložit příslušné DTD
- zajistit kódování stránky na začátek dokumentu i v meta hlavičce dokumentu

#### Literatura:

- [1] Musciano, Ch., Kennedy, B.: HTML a XHTML – kompletní průvodce, Computer Press, Praha 2000
- [2] Pexa, P.“ Tvorba WWW a WAP, Kopp nakladatelství, České Budějovice 2001
- [3] Kosek, J.: HTML – tvorba dokonalých WWW stránek, Grada Publishing, Praha 1998