

Jazyky XHTML, DHTML, CSS a WML

Dodatky

Petr Pexa © 2007
pepe@pf.jcu.cz

1. Upravená kapitola 5.1 Vývojové verze jazyka XHTML

Modularizace XHTML

Modularizace je proces zařazení jednotlivých elementů do fiktivních modulů podle jejich funkce v souvislosti s trendem prohlížení obsahu internetu pomocí jiných zařízení než tradičního PC (mobilní zařízení, tiskárny, čtečky apod.).

- **strukturální modul** – kolekce elementů definujících základní strukturu xhtml dokumentu: `body`, `head`, `html`, `title`
- **textový modul**: – definuje základní elementy sloužící jako kontejner pro text a elementy zaměřené na obsah: `abbr`, `acronym`, `address`, `blockquote`, `br`, `cite`, `code`, `dfn`, `div`, `em`, `h1` až `h6`, `kbd`, `p`, `pre`, `q`, `samp`, `span`, `strong`, `var`
- **bidirectional textový modul**: `bdo`
- **prezentační modul** – obsahuje elementy zaměřené na vzhled, které byly původně v XHTML 1.0 Strict, resp. XHTML1.1, zakázány (ale ostatní verze XHTML 1.0 je povolovaly), nakonec zůstaly tyto: `b`, `i`, `big`, `hr`, `sub`, `sup`, `tt`, `small`
- **hypertextový modul** – modul s jediným elementem: `a`
- **modul seznamů**: `dl`, `dt`, `dd`, `ol`, `ul`, `li`
- **formulářový modul**: `form`, `input`, `label`, `select`, `option`, `textarea`, `button`, `fieldset`, `legend`, `optgroup`
- **modul tabulek**: `caption`, `table`, `td`, `th`, `tr`, `col`, `colgroup`, `tbody`, `tfoot`, `thead`
- **modul obrázků**: `img`
- **module obrazových map**: `area`, `map`
- **modul skriptů**: `noscript`, `script`
- **modul stylů**: `element style`, `atribut style`
- **objektový modul**: `object`, `param`
- **metainformační modul**: `meta`
- **modul linků**: `link`
- **base modul**: `base`

XHTML 1.1 - modulově založený XHTML

Rozsáhlá sada modulů pro komplexnější tvorbu XHTML dokumentů, vynechává již prakticky všechny prezentační vlastnosti (a také rámy) a formátování webu je nutné provést **pouze pomocí CSS**. XHTML 1.1. je tedy velice podobný XHTML 1.0 Strict (viz dále). V XHTML 1.1 dokumentech je dále nutné použít v hlavičce speciální MIME-type `application/xhtml+xml`, který ale podporují jen moderní browsery (vyjma Exploreru) a mobilní telefony s WAP 2.

```
<meta http-equiv="content-type"
      content="application/xhtml+xml; charset=utf-8" />
```

Poznámka: seznam povolených elementů v jazyce XHTML 1.1 najdete v předchozím odstavci, nebo na www.w3.org/TR/xhtml11/doctype.html#s_doctype/

2. Upravená kapitola 5.3 XML deklarace

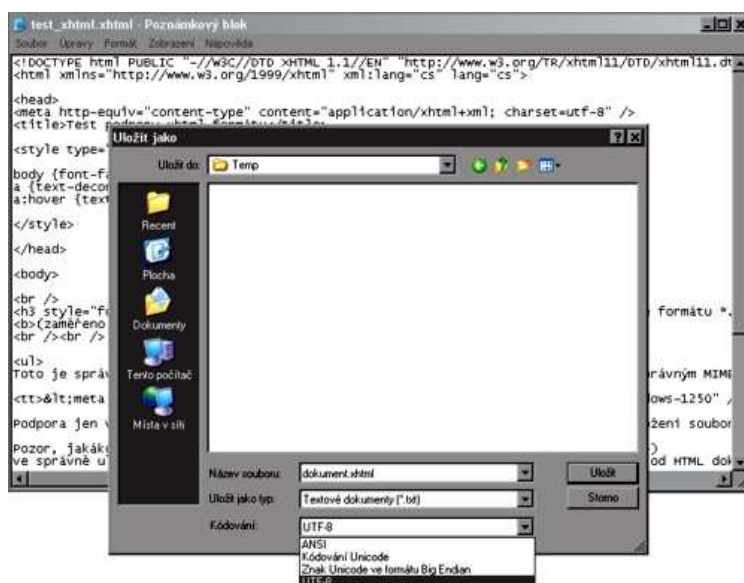
Každý správně formátovaný XHTML dokument by měl na začátku zdrojového kódu obsahovat informaci o použitém kódování a verzi jazyka XML.

```
<?xml version="1.0" encoding="windows-1250"?>
<?xml version="1.0" encoding="iso-8859-2"?>
<?xml version="1.0" encoding="utf-8"?>
```

Pokud ale použijeme v hlavičce stránky kódování **utf-8** pomocí elementu `<meta>` (místo tradičního **windows-1250**), není tento řádek povinný (v Internet Exploreru 6 totiž způsobuje přepnutí do nestandardního zobrazovacího quirk módu).

```
<meta http-equiv="content-type" content="text/html;
      charset=utf-8" />
```

Použitý HTML editor ale musí uložení dokumentu v tomto kódování podporovat! Pozor tedy např. na editor Golden, který ukládá vždy do windows-1250, doporučuji např. PSPad, AceHTML, je možné použít i Poznámkový blok (viz následující obrázek).



Nestačí tedy nastavit kódování utf-8 pouze v hlavičce a dokument pak uložit ve windows-1250, místo českých znaků jsou pak v prohlížeči zobrazeny nesmysly, pokud je použita čeština i v titulku stránky, v Exploreru se dokonce zobrazí jen prázdná stránka!

Poznámka: k tomuto tématu tedy ještě doporučuji profesionálům prostudovat problematiku tzv. zobrazovacích módů prohlížečů, především nestandardního '**quirk módu**' MS Internet Exploreru (např. na <http://www.jakpsatweb.cz/css/mody-prohlizecu.htm>, <http://interval.cz/podklady/bien/730/rezimy.htm> a <http://wellstyled.com/singlelang.php?lang=cs&page=html-doctype-and-browser-mode.html>).

3. Upravená kapitola 5.4 DTD (Document Type Definition)

Každý správně formátovaný XHTML dokument by měl na začátku obsahovat informace o verzi použitého XHTML a typu DTD (Document Type Definition). DTD je jinými slovy návod pro prohlížeč, jak má dokument zpracovat, říká mu, jaké elementy dokument používá a jak s nimi má zacházet. Můžeme použít veřejné DTD typu PUBLIC (viz dále) a lokální DTD pro konkrétní dokument typu SYSTEM (použitelné především u XML dokumentů). Element DOCTYPE není součástí XHTML dokumentu, není elementem XHTML a nemusí mít koncovou značku.

Striktní DTD: DTD XHTML 1.0 Strict

Strict (striktní) definice je určena pro stránky, jejichž autoři se rozhodli zbavit je všech pozůstatků dob minulých. Takové stránky již neobsahují žádné nedoporučované elementy ani atributy jazyka XHTML. Chybí elementy a atributy, které řídí prezentaci a vzhled dokumentu (jako například element <center> či) a u některých elementů není podporován např. atribut *align*. Vzhled stránky je tedy upraven **pouze pomocí kaskádových stylů** (CSS – Cascading Style Sheets). Seznam povolených elementů najdete v předchozí kapitole Vývojové verze XHTML - Modularizace XHTML.

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"
    "http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
```

Html je název kořenového elementu, *PUBLIC* určuje použití veřejného DTD na uvedené url.

Přechodové DTD: DTD XHTML 1.0 Transitional

Druhým typem je tzv. přechodová definice DTD. Ta se nejvíce přibližuje současnému standardu jazyka XHTML. Stránky s tímto typem DTD jsou velmi běžné, neboť používají vše, na co jsme z HTML zvyklí, tady např. všechny elementy i atributy jazyka HTML 4.01. Nepoužívají se však rámy (frames).

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN"
    "http://www.w3.org/tr/xhtml1/DTD/xhtml1-transitional.dtd">
```

DTD s podporou rámců: DTD XHTML 1.0 Frameset

Třetí definice DTD slouží pro tvorbu rámců (frames). Je totožná s přechodovou definicí DTD, jediným rozdílem je, že v této definici nalezneme místo elementu <body> elementy <frameset> a <frame>. V XHTML 1.0 Strict (a XHTML 1.1) nejsou rámy povoleny vůbec.

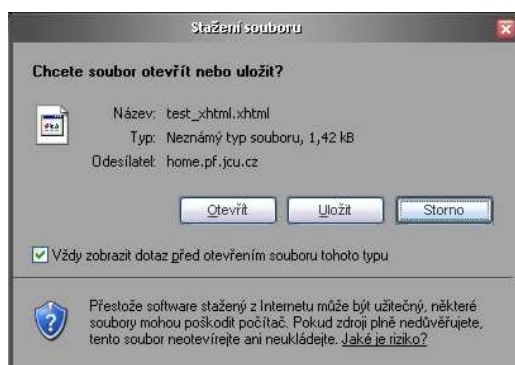
```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Frameset//EN"
    "http://www.w3.org/tr/xhtml1/DTD/xhtml1-frameset.dtd">
```

DTD pro webové dokumenty, napsané v jazyce XHTML 1.1

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.1//EN"
"http://www.w3.org/TR/xhtml11/DTD/xhtml11.dtd">
```

V XHTML 1.1 dokumentech je navíc nutné použít v hlavičce speciální MIME-type `application/xhtml+xml`, nelze použít tradiční `text/html`. Dokumenty napsané v této verzi jazyka XHTML tedy musí být uloženy přímo ve formátu XHTML (s příponou `*.xhtml`), ten ale zatím nepodporuje Explorer ani ve verzi 7 a je nabídnuto pouze stažení a uložení souboru (viz následující obrázek). Také je nutné ověřit, zda hostingový server tento MIME type podporuje. V Opeře a Firefoxu je již tento nový typ dokumentů zobrazen dobře.

```
<meta http-equiv="content-type"
content="application/xhtml+xml; charset=utf-8" />
```



S tímto formátem je ale spojena poměrně revoluční změna v interpretaci kódu prohlížečem - jakákoliv syntaktická chyba (např. `<br>` místo `<br />`) ve správně uloženém XHTML dokumentu již není prohlížeči tolerována (na rozdíl od HTML dokumentů), je hlášena chyba v kódu a stránka se vůbec nezobrazí (viz následující obrázek s chybovým hlášením Opery 9).



DTD typu SYSTEM pro dokumenty, napsané přímo v jazyce XML

```
<!DOCTYPE cenik SYSTEM "cenik.dtd">
```

Cenik je název kořenového elementu, *SYSTEM* určuje použití lokálního DTD *cenik.dtd* s definicí typu dokumentu. Nesporná výhoda tohoto řešení je možnost použití jednoho DTD na více dokumentů:

```
<?xml version="1.0" encoding="windows-1250"?>
<!DOCTYPE cenik SYSTEM "cenik.dtd">
<cenik>
...
</cenik>
```

V úvahu přichází i možnost uložení DTD přímo do XML dokumentu:

```
<?xml version="1.0" encoding="windows-1250"?>
<!DOCTYPE cenik [
... DTD ...
]>

<cenik>
...
</cenik>
```

4. SEO (Search Engine Optimization – optimalizace webové stránky pro vyhledávače), nová kapitola

ZÁKLADNÍ OPTIMALIZAČNÍ PRVKY WEBU (převzato z www.jakpsatweb.cz)

Co vyhledávače sledují	Syntaxe XHTML	Důležitost
Titulek	<title>text titulku</title>	obrovská
Klíčová slova	<meta name="keywords" content="slovo, slovo" />	značná, někdy již žádná (Google)
Popis (description)	<meta name="description" content="Stručný popis" />	různá
Nadpis 1. úrovně	Nadpis	značná
Ostatní nadpisy	Nadpis	sporná
Začátek stránky	<body>Několik prvních slov ...	větší než malá
Adresa URL	jméno souboru včetně cesty	různá
Text odkazů mířících na stránku z jiného serveru	Text odkazu	u některých vyhledávačů obrovská (Google, Jyxo)
Titulky u odkazů	odkaz	značná
Alty u obrázků		malá
Text stránky	text	malá
Katalogový popis	co zadáte do katalogu	různá (velká na Seznamu)

DALŠÍ OPTIMALIZAČNÍ PRVKY WEBU

- **Vytvoření tzv. zpětných odkazů** – u nových webů registrace do katalogů (Seznam, Centrum, Atlas, apod.)
- Doporučuje se mít na stránce alespoň **jednou klíčové slovo tučně a jednou kurzívou**, význam je malý.
- **Roboti velmi často neindexují Javascript**. Pokud ho použijeme např. pro navigaci (menu), je vhodné ho „schovat“ do souboru (<script type="text/javascript" src="x.js"></script>).
- Pouze **Flash** animaci není vhodné použít pro navigaci, protože robot vyhledávače zaindexuje pouze první stranu a ostatní stránky již nevidí. Pokud je na stránce Flash, je nutné ho doplnit textovými informacemi, popřípadě udělat také neflashovou verzi.
- **Klikací obrázkové mapy** - ne všechny vyhledávače si s tímto typem navigace dokáží poradit, vhodné i textové odkazy.

- **Splash page** (titulní stránka, která obsahuje pouze veliký obrázek, klikací mapu nebo flashovou animaci), vhodné doplnit i textovým odkazem (tzv. Skip Intro).
- **Pro roboty jsou velký problém rámce (rámy, frames)**, odmítají je indexovat a zaindexují pouze část noframes. Pokud robot zaindexuje obsah jednotlivých stránek v rámcích, vyhledavače nabídnou odkaz na konkrétní stránku a ne rámcový index, takže se většinou zobrazí stránka bez navigačního menu.
- **Dynamická URL** obsahuje otazník, za kterým si script předává parametry. Například www.example.com/clanek.php?clanek=1765. Vyhledavače většinou neindexují stránky s mnoha parametry. Například Google indexuje maximálně tři proměnné, ale některé menší vyhledavače to nedokáží vůbec. Pro uživatele, ale i pro vyhledavače je jednodušší zapamatovat si statickou URL stránky. Příkladem takového URL může být www.example.com/nadpis-clanku.html.
Viz problematika tzv. **mod_rewrite** na www.pf.jcu.cz/pepe/Diplomky/vacek.pdf.

Užitečné webové adresy ke studiu problematiky SEO:

<http://www.jakpsatweb.cz/vyhledavace.html>
<http://www.jakpsatweb.cz/katalog/seo.html>
<http://vyhledavace.info/seo-faq/>
<http://seo.nawebu.cz/>

5. Doplnění kapitoly 14 - Fotografie a grafika

- ***.raw** - formát obsahuje nekomprimovaná data přímo ze senzoru fotoaparátu (tzv. digitální negativ), profesionální úpravy jednotlivých RGB kanálů, až 8x větší objem dat

6. Doplnění kapitoly 19 - Zvuk na webové stránce

ZVUKOVÉ FORMÁTY PRO WEB

- ***.wav** - nekomprimovaný záznam reálného zvuku, pro web tedy méně vhodný
- ***.mid** - kódovaná imitace zvuku hudebních nástrojů, zvuk generován např. zvukovou kartou či syntezátorem (samplerem), malá velikost souborů, nutný kvalitní HW
- ***.mp3** - komprese MPEG1-Layer 3 (nejčastěji 1:12), volitelný bitrate (datový tok, např. 128 kb/s pro CD kvalitu), textový popis mp3 souborů (tzv. ID3 Tagy)
- ***.wma** - komprimovaný formát zvuku firmy Microsoft, podobný mp3
- ***.ogg-vorbis** - další ztrátově komprimovaný formát zvuku
- ***.rm** - RealMedia, technologie firmy RealNetworks pro streamované vysílání zvuku

SW pro záznam, tvorbu, zpracování a úpravu zvukových souborů:

Audiograbber (wav), Lame Encoder (mp3), Cubase (midi), Audacity, Nero Wave Editor, Acoustica, Easy Audio Editor, atd.

7. Doplnění kapitoly 20 - Video na webové stránce

VIDEOFORMÁTY PRO WEB

- ***.avi** - komprimovaný i nekomprimovaný základní záznam obrazu i zvuku původně pro Windows, pro web vhodný jen s kompresí (MPEG-4, kodeky - např. DivX, XviD)
- ***.mpg, mpeg, mp4** - záznam obrazu s vysokou kompresí, používaný např. na DVD
- ***.wmv** - komprimovaný formát obrazu firmy Microsoft, podobný mpg
- ***.3gp** - další komprimovaný formát obrazu, používaný především v mobilních telefonech
- ***.mov** - multimediální vícestopý formát QuickTime firmy Apple, vhodný zejména pro internetové vysílání (průběžné zobrazování - streaming, webcasting)
- ***.rm** - RealMedia, technologie firmy RealNetworks pro streamované vysílání videa
- ***.asf** - konkurenční technologie firmy Microsoft pro streamované vysílání videa

PŘEHLED KOMPRESNÍCH STANDARDŮ A KODEKŮ

MPEG-1: kompresní standard (1:12) digitálního videa (VideoCD) a zvuku (MP3), rozlišení 352x288px, rychlost přenosu jen 1,5 Mb/s, pro web tedy méně vhodný

MPEG-2: lepší metoda prokládání rádků, rozlišení 352x288px při 4 Mb/s (S-VHS); 720x576px, 10Mb/s (DVD); 1440x1440px, 30Mb/s (HDTV) a 1920x1080px, 40Mb/s (HD DVD, Blu-ray Disc)

MPEG-4: velká komprese videa při srovnatelné kvalitě s originálem, variabilní bitrate až 6MB/s (např. s použitím kodeku DivX).

DivX: obrazový kodek pro kompresi videa kompatibilní s MPEG-4, 7GB dat → 700MB

XviD: open source verze kodeku DivX

SW pro záznam, zpracování, úpravu a kompresi videosouborů:

Pinnacle Studio, Windows Movie Maker, Virtual Dub (DivX, XviD komprese), Windows Media Encoder (streamované vysílání videa, webcasting)