



# HTML

## HyperText Markup Language

# HTML – historie, verze

- Hypertext
- Jazyk popisující strukturu dokumentu
- SGML/XML Aplikace
- Verze:
  - HTML 1 – 1990+ TBL, CERN
  - HTML 2.0 – 1995 as RFC 1866
    - kodifikace (zachycení a standardizace) aktuálního stavu jazyka
    - všechny základní elementy (P, UL, PRE, FORM, ...)
  - HTML 3.0 – 1995, proposed W3C standard
    - pokus o silný standard, nepoužívané
    - obsahovalo mj. matematické vyznačování

# HTML – verze

- HTML 3.2 – 1997, W3C recommendation
  - kodifikace (zachycení a standardizace) aktuálního stavu jazyka
  - nové elementy: TABLE, DIV, FONT, MAP, APPLET, etc.
- HTML 4.0 – 1998, W3C recommendation
  - formálně silný základ, praktické použití; důraz na přenositelnost, přístupnost
  - nové elementy a atributy: STYLE, FRAME, OBJECT, SCRIPT, lang, class, accesskey, etc.
  - vylepšení: TABLE, FORM
- HTML 5 – 2013, Candidate recommendation

# XHTML

- SGML → XML
  - zjednodušení DTD
  - snazší strojové zpracování, výměna dat
  - lepší modularita a rozšiřitelnost jazyka
- XHTML 1.0 – 2000, W3C recommendation
  - HTML 4.01 jako XML aplikace
  - nasměrování k čistému logickému vyznačování
- XHTML 1.1 – 2001, W3C recommendation
  - modularizace XHTML1
- XHTML 2 – dosud jako working draft
  - cíl: obecnější textové vyznačování, zcela bez prezentačních prvků

# Text v HTML

`<h1>HyperText Markup Language</h1>`

`<p>HTML is the lingua franca for publishing hypertext on the <abbr title="World Wide Web">WWW</abbr>.`

`<a href="http://www.w3.org/TR/REChtml40/">HTML 4.0</a> is W3C's recommendation for the latest version of HTML.</p>`

# Obecné prvky

- Značky: vyznačují elementy obsahu
  - <znacka> obsah </znacka>
  - atributy (u otevírací značky)
  - velikost písmen
  - prázdné elementy
  - img, br – nepárové značky
  - h1 není příkaz, href není parametr
- Komentáře
  - <!-- komentář -->
- Entity
  - &lt; &gt; & &
  - &acute; é &#x41;A
- Bílé místo
  - &nbsp;

# HTML x XHTML

## HTML = aplikace SGML

- Značky
  - case insensitive
  - možno vynechat uzavírací
- Atributy
  - atribut=hodnota
  - atribut="hodnota s mezerou"
  - atribut
- Ne-SGML data
  - `<![CDATA[ ... ]]>`
  - obvykle stačí komentáře
- Renderování
  - volná interpretace, tolerance

## XHTML = aplikace XML

- Značky
  - case sensitive: malými
  - uzavírací povinně  
(`<p>... </p> <img ... />`)
- Atributy
  - povinné uvozovky
  - žádná minimalizace
- Ne-XML data
  - povinně CDATA sekce
  - styly, JavaScript atd lépe do externích souborů
- Renderování
  - striktní chování

# Varianty: Strict × Transitional

- Strict
  - pouze logické vyznačování
- Transitional
  - pojem „deprecated element“ (celkem 10 v HTML4)
  - Nestandardní rozšíření
    - staré verze HTML (Netscape, Microsoft, ...)
  - Důsledky
    - sada elementů, struktura těla dokumentu
    - chování prohlížečů (CSS)
- Strict je důležitější než XHTML



# HTML dokument

- Preambule, deklarace
  - Xml deklarace
  - DOCTYPE
- Záhlaví - <head>
  - title, meta, style...
- Tělo

# Příklad HTML dokumentu

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"
"http://w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
<html xmlns="http://w3.org/1999/xhtml" lang="en">
<head>
<title>HTML Home Page</title>
<meta name="keywords" content="HTML, XHTML 1.0" />
<style type="text/css"><!--
body { margin-left: 10%; }
// -->
</style>
</head>
<body>
<h1 class="c1">HyperText Markup Language</h1>
<p>HTML is the <i>lingua franca</i> for publishing hypertext on the
<abbr title="World Wide Web">WWW</abbr>. <a
href="http://www.w3.org/TR/REC-html40/">HTML 4.0</a> is W3C's
recommendation for the latest version of HTML.</p>
</body>
</html>
```

# Preamble, deklarace

- Preamble
    - SGML: implicitní
    - XML: povinná
      - `<?xml version="1.0" encoding="UTF-8"?>`
  - Deklarace typu dokumentu
    - `<!DOCTYPE kořenový_element SYSTEM/PUBLIC "URL">`
    - označuje gramatiku
    - **!!! povinná (interpretace dokumentu)**
    - **HTML4 strict:**
      - `<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN"`
      - `"http://www.w3.org/TR/html4/strict.dtd">`
    - **HTML4 Transitional**
      - `... "-//W3C//DTD HTML 4.01 Transitional//EN"`
      - `"http://www.w3.org/TR/1999/REC-html401-19991224/loose.dtd"`
    - **XHTML Strict**
      - `... "DTD XHTML 1.0 Strict//EN"`
      - `"http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd"`
- + XML deklarace

# Záhlaví <head>

- Meta-informace o dokumentu
  - není obsah dokumentu, nezobrazuje se
- Dokument element
  - HTML: atributy xmlns, lang, ...
- Značky záhlaví
  - TITLE
  - META: metainfo (author, keywords) + HTTP
    - vyhledávače, filtrování obsahu, ...

```
<meta http-equiv="Content-Type" content="text/html; charset=windows-1250"/>
```
  - LINK: odkazy na úrovni dokumentů
    - typy: next, prev, contents, index, copyright, ... , stylesheet + alternate
  - STYLE
    - type, media
  - SCRIPT, NOSCRIPT
    - type, src

# Tělo <body>

- **Obsah dokumentu**
  - prohlížeč: canvas pro zobrazení
  - BODY
  - FRAMESET
- **Atributy**
  - style, class, id, title, lang
- **Vnořené elementy**
  - blokové, v nich textové (inline)
  - struktura rámců

# Obsahové elementy

- **Blokové**
  - zalamují odstavec; hierarchie
  - bloky, tabulky, formuláře
- **Textové**
  - uvnitř blokových
  - frázové × prezentační
- **Generické**
  - kontejnery, vazba na CSS
- **Obecné atributy**
  - všechny elementy
  - id, class, style, title ; lang, dir ; onSomeEvent

# Základní blokové elementy

- Odstavec – `p`
- Nadpisy – `h1-h6`
- Odrážky, číslování – `ul/ol li, dl dt dd`
- Prezentační atributy – `align=center`
- Uvozovkování – `blockquote` – atribut `cite="URI"`
- Předformátovaný text – `pre`
- Odřádkování – `br`

# Základní textové elementy

- Důraz – `em`
- Zesílení – `strong`
- Podtržení, škrtnutí – `ins`, `del`
- Tučné, kurzíva, podržení – `b`, `i`, `u`
- Indexy – `sub`, `sup`
- Monospaced text – `tt`
- Další – `cite`, `abbr`, `q`, `code`
- Zapomeňte na:
  - `strike`, `big/small`, `font`
  - `<xyz align="left" background="..." > a`  
podobné



# Hypermediální elementy

- **Odkaz** – `a (name, href, title)`
- **Obrázek** – `img;`
- **Klikací mapy**
  - `map, area, src, alt, align, usemap`
  - `name, shape, coords, href`
- **Další** – `object, applet, param`

# Generické kontejnery

- **Bez formátování**
- `div`
  - **blokový**
- `span`
  - **Řádkový (inline)**
- **Atributy pro stylování**
  - `id, class, style`

# Formulářové elementy

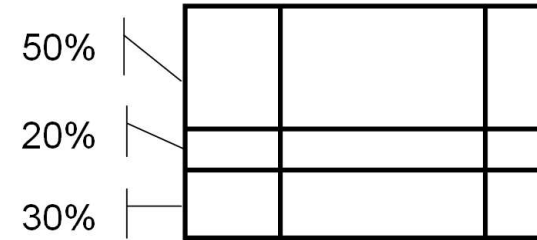
- Interakce klient (uživatel) – server
  - zasílání dat na server, zpracování URI objektem
  - metody GET × POST
- form – method, action, enctype
- input – name, value; type; size, maxlength, checked
  - type:  
text | password | checkbox | submit | file | hidden
- select – multiple, option selected, optgroup
- textarea – rows, cols
- label for; fieldset, legend
- **Obecné atributy**
  - tabindex, accesskey, disabled, readonly

# Tabulky

- Popis dat s tabelární strukturou
  - zneužití: formátování pro HTML 3
- `table; caption`
  - **prezentační atributy:** `width, border`
- `tr; th, td`
  - `colspan, rowspan`
  - **prezentační atributy:** `align, valign`
- `thead, tfoot, tbody`
  - kontejnery řádků
- `col, colgroup`
  - stylování sloupců

# Rámce

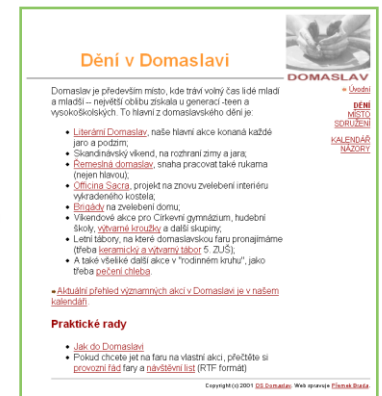
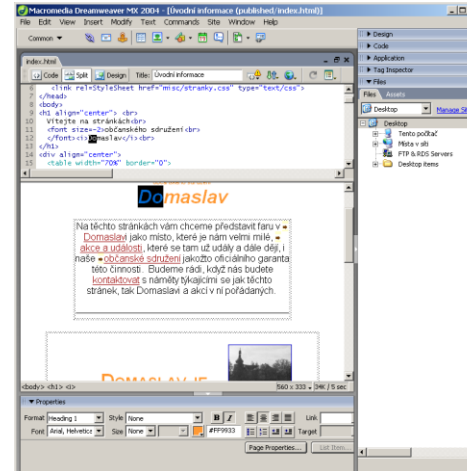
```
<FRAMESET cols="150,4*,1*"
rows="50%,20%,30%" >
```



- frameset: definice mřížky
  - title; rows, cols – velikosti rámců
  - absolutní: pixely (“30”)
  - relativní: procenta (“25%”),
  - poměry (“4\*”)
- noframes
  - alternativní obsah pro non-frame prohlížeče
- frame: úvodní obsah rámce
  - name, src
  - noresize, scrolling = “auto|yes|no”, frameborder
- <A href=“...”> s rámci
  - target=“#name”; \_blank, \_self, \_parent, \_top

# Tvorba správného HTML

- Editory
  - textové
  - „značkovací“
  - WYSIWYG
- Generování
  - z dokumentů
  - aplikacemi
- Další nástroje
  - Validátor
  - <http://browsershots.org/>
- Zobrazení
  - zdrojový kód
  - prohlížeče – textové, grafické, čtečky, jiné (Fangs)



# Strukturování obsahu

- Varianta „prezentační“
  - HTML 3 bez CSS
  - vyznačit tak, aby se co nejlépe zobrazilo
  - tabulkový layout
- Varianta „informační“
  - HTML4/XHTML s CSS
  - vyznačit tak, aby se co nejlépe vyhledávalo
  - důležitý obsah napřed
  - `dobrý title a h1/h2`

# Přístupnost

- Přístupnost = bezbariérovost
  - prohlížeč, OS, rozlišení, scriptování
  - kultura, motorické schopnosti, vidění
  - extrémně handicapovanými návštěvníky jsou vyhledávací roboti
- Zásady
  - validovat
  - používat informační strukturování
  - title, hierarchie nadpisů, oddělená navigace, linearizace tabulek
  - psát stručně a srozumitelně
  - členění textu
  - zpřístupnit formuláře
  - vyhnout se rámcům



# Možnosti HTML pro přístupnost

- **Elementy**

- `h*, p; div`
- `em, strong, q, cite, ...`
- `fieldset, legend; optgroup; label` for
- `th, thead/tfoot; caption`

- **Atributy**

- `title, lang, dir, accesskey`
- **specifické**
  - `table - longdesc`
  - `input - title, tabindex;`
  - `img - alt, title, longdesc; a title`

# Tipy

```
<p>  
  Ceník najdete  
  <a href="cenik.html">zde</a>.  
</p>
```

```
<a href="ut/p/c5/">  
    
</a>
```

```
<div class="container">  
  <div class="head">  
    My Heading  
  </div>
```

```
  <div class="body">  
    <div class="text">  
      Text goes here  
    </div>  
  </div>  
</div>
```

```
<p>  
  Prohlédněte si náš  
  <a href="cenik.html">ceník</a>.  
</p>
```

```
<a href="ut/p/c5/" title="...">  
    
</a>
```

```
<div class="container">  
  <h1>My Heading</h1>  
  <p>My Paragraph</p>  
</div>
```