

# INTERNETOVÉ TECHNOLOGIE PRO E-BUSINESS C1

## 2008-10-09

Jan Tyrychtr

[tyrychtr@pef.czu.cz](mailto:tyrychtr@pef.czu.cz)

PEF, 305? nebo rektorát u Šimka

### ÚVOD DO HTML, XML A XHTML:

#### HTML:

##### ✓ HyperText Markup Language:

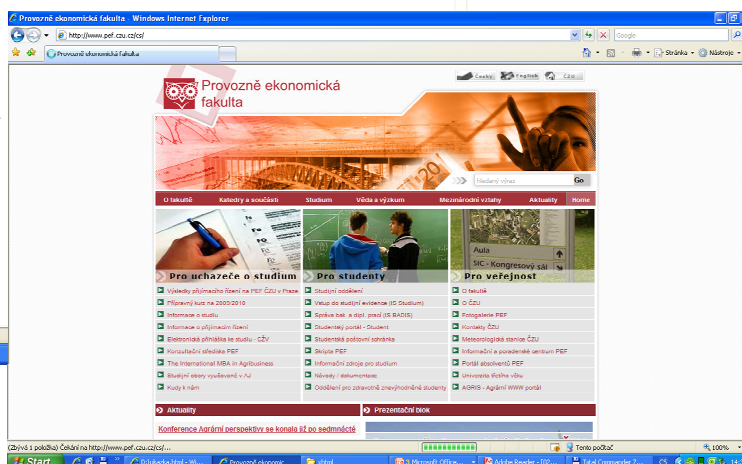
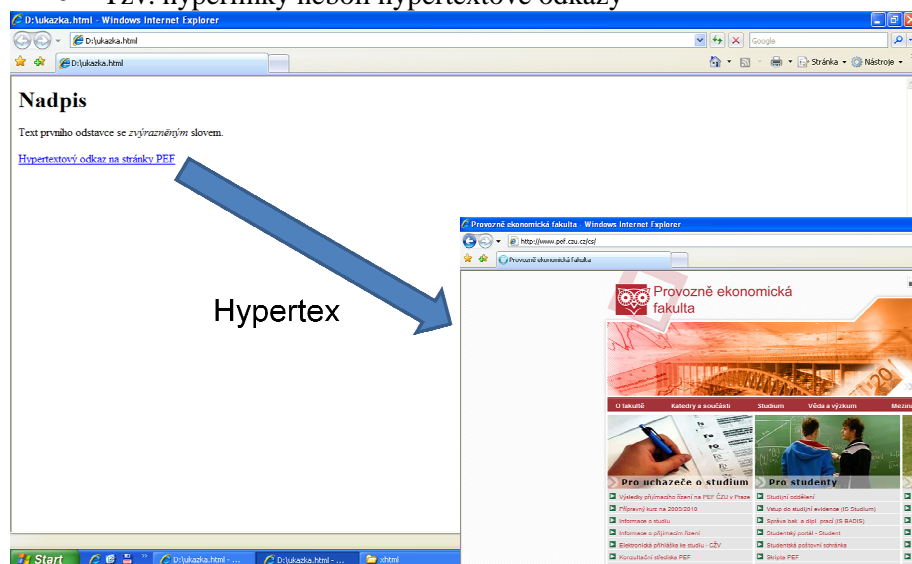
- Značkovací jazyk pro hypertext
- Jazyk pro vytváření stránek WWW

##### ✓ Co je **značkovací jazyk**?

- Jazyk, jehož zdrojový text obsahuje současně jak vlastní text, tak instrukce pro jeho zpracování.
- V HTML instrukce představují značky (tagy)

##### ✓ Co je **hypertext**?

- Zobrazuje informace v textu, které odkazují na upřesnění
- Tzv. hyperlinky neboli hypertextové odkazy



##### ✓ Jak si to celé představit? **Značkovací jazyk s textem:**

```
<h1>Nadpis</h1>
```

```
<p>Text prvního odstavce se <em>zvýrazněným</em> slovem.</p>
```

```
<a href="http://www.pef.czu.cz"> Hypertextový odkaz na stránky PEF</a>
```

#### HTML – 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, KBD, FORM, ...)

##### ✓ **HTML 3.0** – 1995, proposed W3C standard

- Pokus o silný standard, nepoužívané
- Obsahovalo mj. matematické vyznačování

##### ✓ **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
  - Vylepšení: TABLE, FORM

### HTML – obecné prvky:

- ✓ **Značky** – vyznačují elementy obsahu  
<znacka> obsah </znacka>  
„Nepárové značky“: <img>, <br>
- ✓ **Atributy** – popisují vlastnosti elementu, u otevírací značky  
<znacka atribut=„hodnota“>obsah</znacka>
- ✓ **Komentáře:**  
<!-- komentář -->
- ✓ **Entity:**  
    &lt;         <                     &gt;         >  
    &amp;       &                    &eacute;       é  
    &#x41;     A                    &nbsp;         bílé místo (prázdná mezera)

### HTML dokument:

- ✓ **1.:**  

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN"
"http://www.w3.org/TR/html4/loose.dtd">
```

Preamble,  
deklarace

```
<html>
```

Záhlaví

```
<head>
```

```
</head>
```

```
<body>
```

Tělo

```
</body>
```

```
</html>
```

- ✓ **2.:**  

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN"
"http://www.w3.org/TR/html4/loose.dtd">
```

```
<html>
```

```
<head>
```

```
<meta http-equiv="content-type" content="text/html; charset=iso-8859-1" />
```

```
<meta http-equiv="Description" content="Stránka slouží k výuce základů
HTML" />
```

```
<meta name="author" content="Tygr" />
```

```
<title>Výuka HTML</title>
```

```
</head>
```

```
<body>
```

```
<h1>Nadpis</h1>
```

```
<p>Text prvního odstavce se <em>zvýrazněným</em> slovem.</p>
```

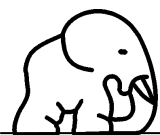
```
<a href="http://www.pef.czu.cz"> Hypertextový odkaz na stránky PEF</a>
```

```
</body>
```

```
</html>
```

### DTD:

- ✓ **DTD** (Document Type Definition, česky Definice typu dokumentu) – je jazyk pro popis struktury XML případně SGML dokumentu
  - Vymezuje jazyky HTML a XHTML
  - Popisuje, jak mohou být značky navzájem uspořádány a vnořeny



- Vymezuje atributy pro každou značku a typ těchto atributů
- ✓ **Nevýhody:**
  - DTD je poměrně starý a není v podobě XML.
  - Silnější jazyk pro popis XML dokumentů je např. XML Schema

✓ **Zápis v XHTML:**

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

**Validní kód:**

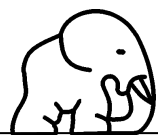
- ✓ Zda dokument splňuje pravidla daná svou definicí typu dokumentu, se vyjadřuje slovním spojením validita dokumentu.
- ✓ Pokud je v DTD například definováno, že element <a> nesmí obsahovat další elementy <a> a v kódu se vyskytne například <a href="URL">odkaz<a href="URL">odkaz2</a></a>, dokument je not valid („nevalidní, neplatný“).
- ✓ Validátor naleznete na <http://validator.w3.org/>

**XML:**

- ✓ **SGML** (Standard Generalized Markup Language), vznik 1986, norma ISO 8879. Obecný značkovací jazyk umožňující definici vlastních (podřízených) značkovacích jazyků na základě vlastní definice typu dokumentu – DTD. HTML je jednou z aplikací SGML, každá verze HTML používá značky tak, jak je popsáno v každé příslušné DTD.
- ✓ **XML** (Extensible Markup Language)
  - Jde o zjednodušenou verzi toho nejdůležitějšího ze SGML.
  - Je v něm pevně dáno několik parametrů, jako používané oddělovače nebo maximální délka názvů značek.
  - Na rozdíl od HTML není definován způsob zobrazení, ale samotná data. Způsob zobrazení je pak umístěn v jiných souborech.
  - Ze SGML si ale zachovává hlavně možnost definování vlastních DTD.
- ✓ Odtud je už jen krůček k přeformulování **HTML do XML**. Výsledkem tohoto přeformulování je nový jazyk, XHTML 1.0.
- ✓ Aby byl dokument považován za **správně strukturovaný** („well-formed“), musí mít nejméně následující vlastnosti:
  - Musí mít právě **jeden kořenový** (root) element.
  - Neprázdné elementy musí být ohraničeny **startovací a ukončovací značkou**. Prázdné elementy mohou být označeny tagem „**prázdný element**“.
  - Všechny **hodnoty atributů musí být uzavřeny v uvozovkách** – jednoduchých (') nebo dvojitých ("), ale jednoduchá uvozovka musí být uzavřena jednoduchou a dvojitá dvojitou. Opačný pár uvozovek může být použit uvnitř hodnot.
  - **Elementy mohou být vnořeny, ale nemohou se překrývat**; to znamená, že každý (ne kořenový) element musí být celý obsažen v jiném elementu.
  - **Jména elementů v XML rozlišují malá a velká písmena**: např. „<Příklad>“ a „</Příklad>“ je pár, který vyhovuje správně strukturovanému dokumentu, pár „<Příklad>“ a „</příklad>“ je chybný.

✓ **Ukázka:**

```
<?xml version="1.0" encoding="windows-1250"?>
<seznam predmet="e-business" druh="cvičení" >
<titulek>Seznam studentů </titulek>
  <student cislo= "1">
    <jmeno>Jirka</jmeno>
    <prijmeni>Vágus</prijmeni>
    <obor>INFO</obor>
    <kontakt>
      <email>vagus@email.cz</email>
      <icq>123456</icq>
    </kontakt>
  </student>
</seznam>
```



✓ **Aplikace:**

- **XHTML** – Nástupce jazyka HTML.
- **RDF** – Resource Description Framework, specifikace, která umožňuje popsat metadata, např. obsah a anotace HTML stránek.
- **RSS** – je rodina XML formátů, sloužící pro čtení novinek na webových stránkách
- **SMIL** – Synchronized Multimedia Integration Language, popisuje multimedia pomocí XML.
- **MathML** – Mathematical Markup Language je značkovací jazyk pro popis matematických vzorců a symbolů pro použití na webu.
- **SVG** – Scalable Vector Graphics je jazyk pro popis dvourozměrné vektorové grafiky, statické i dynamické (animace).
- **DocBook** – Sada definic dokumentů a stylů pro publikační činnost
- **Jabber** – Protokol pro Instant messaging
- **SOAP** – Protokol pro komunikaci mezi Webovými službami
- **Office Open XML, OpenDocument** – Souborový formát určený pro ukládání a výměnu dokumentů vytvořených kancelářskými aplikacemi

**XHTML:**

✓ **XHTML** (extensible hypertext markup language), rozšiřitelný hypertextový značkovací jazyk

✓ Podstatné **změny v XHTML**:

- Dokumenty musí být well-formed
- Tedy například `<p>Toto je <em>zvýraznění</p></em>` je špatně, dochází ke křížení elementů
- Elementy a atributy musí být malými písmeny (lowercase)
- Každý element musí být uzavřený. Tedy `<p>první odstavec <p>druhý odstavec` je špatně. Taktéž nepárové elementy musí být uzavřené, tedy místo `<br>` musí být buďto `<br></br>` anebo lépe zkrácená verze: `<br />`
- Hodnoty atributů musí být vždy v uvozovkách, tedy ne `<td rowspan=4>`, ale `<td rowspan="4">`

✓ **Verze:**

- **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  
HTML = aplikace SGML  
XHTML = aplikace XML

✓ **Moduly:**

- **Strukturální modul** zahrnující prvky, které tvoří základní strukturu XHTML dokumentu (body, head, html, title),
- **Textový modul**, jenž definuje základní prvky sloužící k označení textu a obsahu dokumentů (h1 až h6, address, blockquote, div, p, pre, abbr, acronym, br, cite, code, dfn, em, kbd, q, samp, span, strong, var),
- **Hypertextový modul** s prvky a sloužícím pro hypertextové odkazy na jiné zdroje,
- **Modul seznamů** obsahující prvky sloužící k vytváření seznamů (dl, dd, dt, ol, ul, li),
- **Formulářový modul** s prvky pro tvorbu formulářů (form, input, label, select, textarea aj.).
- **Modul tabulek**
- **Objektový modul**
- **Modul linků**
- **Metainformační modul**

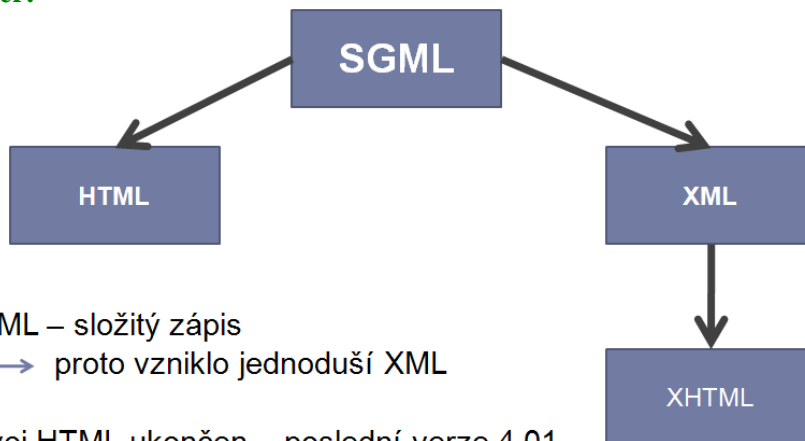
✓ **Specifikace** – existuje ve třech verzích: Strict, Transitional a Frameset:

- **XHTML 1.0 Strict** se používá, pokud chcete oddělit obsah od formy. Předpokládá se jeho užití společně s CSS, které vám umožní dosáhnout potřebných grafických efektů. Nicméně i tato verze obsahuje formátovací elementy, například `<b>` nebo `<i>`.
- **XHTML 1.0 Transitional** je přechodným DTD pro webové stránky, který vám umožní používat překonané tagy.
- **XHTML 1.0 Frameset** vám umožňuje používat zastaralé značky jako XHTML 1.0 Transitional a přidává podporu pro rámce.



- ✓ **Značky**, které byste si měli pamatovat:  
`<html>`, `<head>`, `<title>`, `<body>`  
`<a href=URL>odkaz</a>`  
`<br />` - nový řádek  
`<em>kurzíva</em>`  
`<div>oddíl</div>`  
`<h1>` až `<h6>` - `<h1>Nadpis</h1>`  
`<p>odstavec</p>`  
`<span>úsek textu</span>`  
`<strong>tučný text</strong>`

### Závěr:



SGML – složitý zápis

→ proto vzniklo jednodušší XML

Vývoj HTML ukončen – poslední verze 4.01

→ vychází z něj XHTML 1.0 jako XML aplikace

- ✓ Pro váš e-business nepotřebujete znát všechny XHTML **značky** – pro e-business je důležité znát technologie!
- ✓ Zapamatujte si poslední verze značkovacích jazyků. Jaký je tedy hlavní **rozdíl** mezi x(html) a xml?  
Odpověď: V XML SI ZNAČKY MOHU VYTVÁŘET (např. `<student></student>`,...). V (X)HTML MOHU POUŽÍVAT JEN POVOLENÉ ZNAČKY (např. `<p></p>`, `<div></div>`,...)!

### Zajímavé odkazy:

- ✓ <http://www.w3c.org> resp. <http://www.w3.org>
- ✓ <http://www.jakpsatweb.cz>
- ✓ <http://interval.cz> – webdesign

Příklad:

- ✓ Zkuste si vytvořit vlastní životopis v XHTML dokumentu.
- ✓ V příslušném editoru (stačí poznámkový blok) vytvořte základní strukturu XHTML dokumentu.
- ✓ Dále do dokumentu doplňte obsah, mezi značky `<body></body>`.
- ✓ Soubor uložte jako název.html a otevřete v internetovém prohlížeči.

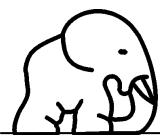
✓ Vzor 1:

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

✓ Vzor 2:

```
<body>
<h1>Životopis</h1>

<p><strong>Jméno a příjmení:</strong> Jan Tyrychtr</p>
```



```
<p><strong>Kontakt:</strong> tyrychtr@pef.czu.cz</p>
<p>
<strong>Vzdělání:</strong><br />
<ul>
  <li><em>2008 - </em> ČZU v Praze, obor Informatika, zakončeno SZZ</li>
  <li><em>2003 - 2006</em> ČZU v Praze, obor Informatika, zakončeno SZB
  </li>
</ul>
</p>
<a href="http://www.pef.czu.cz">Odkaz na stránky PEF</a>
</body>
```