

# PB138 — XHTML

(C) 2019 Masaryk University -- Tomáš Pitner, Luděk Bártek, Adam Rambousek

# XHTML

- Vznikl přepisem HTML pomocí pravidel XML
  - XHTML dokument je dobře utvořeným XML dokumentem.
    - Měl by mít XML prolog.
    - Hodnoty atributů musí být ohraničeny pomocí ", resp. '.
    - Každý element musí mít počáteční a koncovou značku.
    - Elementy se nesmí křížit.
- Aktuální 1.1
- Používaná verze 1.0

## Co je XHTML?

- Rodina současných a budoucích typů dokumentů, které reprodukuje a rozšiřuje HTML 4.
- Tyto typy dokumentů jsou založeny na XML a jsou navrženy tak, aby byly zpracovatelné pomocí uživatelských agentů (prohlížečů, ...) založených na XML.
- XHTML 1.0 je přeformulováním tří HTML4 typů dokumentů jako aplikací standardu XML 1.0.
  - XHTML dokumenty splňují standard XML..
  - Měly by být zpracovatelné stejně dobře, pokud ne lépe v existujících uživatelských agentech podporujících HTML 4 i XHTML 1.0.
- Při dodržení jednoduchých doporučení je zpracovatelný i pomocí uživatelských agentů podporujících HTML 4.0

## Struktura XHTML dokumentu

XHTML 1.0 dokument musí:

- splňovat jedno ze tří DTD.
- Kořenový element musí být element html.
- Kořenový elementu musí:
  - obsahovat deklaraci jmenného prostoru.
  - Jmenný prostor musí být <http://www.w3.org/1999/xhtml>
- Kořenovému elementu musí předcházet deklarace typu dokumentu.
- Ukázka kořenového elementu:

```
<html xmlns="http://www.w3.org/1999/xhtml" xml:lang="en" lang="en">
```

# XHTML Strict

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

- Striktně odděluje strukturní značkování od vzhledu dokumentu.
  - vzhled je nutno řešit pomocí CSS
  - nedefinuje vizuální značky a atributy (např. <b/>, ...)

# XHTML Transitional

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

- Umožňuje kombinovat strukturní a vizuální značkování

# XHTML Frameset

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

- Určeno k vytváření stránek pomocí rámců

# XHTML a jmenné prostory

- XHTML podporuje jmenné prostory + Jmenný prostor pro XHTML:
  - <http://www.w3.org/1999/xhtml>
- Ukázka XHTML 1.0 dokumentu

```
<?xml version="1.0" encoding="UTF-8">
<!DOCTYPE html
  PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"
  "http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
<html xmlns="http://www.w3.org/1999/xhtml" xml:lang="en" lang="en">
  <head>
    <title>Naše první XHTML 1.0 Stránka</title>
  </head>
  <body>
    <h1>První odstavec</h1>
    <p>
      Kdo najde aspoň 5 rozdílů oproti HTML 4?
    </p>
  </body>
</html>
```

## Použití XHTML s jinými jmennými prostory

- Pomocí jmenných prostorů lze XHTML kombinovat s jinými druhy značkování
  - RDF - doplnění sémantiky.
  - SVG - vkládání grafiky
    - chybí podpora v prohlížečích
    - jednodušší vkládání pomocí elementu `img`
  - Voice + XHTML profile - možnost vyplňování formulářů pomocí hlasu
    - nutnost nainstalovat rozšíření
    - funkční pouze v některých prohlížečích
    - lze použít pouze v některých jazycích

## Ukázka použití XHTML s jinými jmennými prostory

- Vkládání jiných značkování do XHTML

```

<?xml version="1.0" encoding="utf-8">
<html
  xmlns="http://www.w3.org/1999/xhtml"
  xml:lang="en" lang="en">
<head>
  <title>Kombinování MathML a XHTML</title>
</head>
<body>
  <p>Následuje ukázka v MathML</p>
  <math xmlns="http://www.w3.org/1998/Math/MathML">
    <apply> <log/>
      <logbase>
        <cn> 3</cn>
      </logbase>
      <ci> x </ci>
    </apply>
  </math>
</body>
</html>

```

## Rozdíly oproti HTML 4

- Dokumenty musí být správně utvořené:
  - ačkoliv například křížení elementů je v SGML zakázáno, tak je v prohlížečích často tolerováno.

```
<p>Ukázka <em>křížení elementů.</p></em>
```

- Elementy a jména atributů musí být malými písmeny.
  - XML je case-sensitive takže, <p> a <P> jsou různé značky.
- Neprázdné elementy musí mít koncovou značku.
- Hodnoty atributů musí být uzavřeny v úvozovkách nebo apostrofech.
- XML (XHTML) nepodporuje minimalizaci atributů:
  - <dl compact="compact"> vs. <dl compact>
- Prázdné elementy musí být ukončeny:
  - nelze např. použít <br>, musí se zapisovat <br/>
- Zpracování bílých znaků v attributech:
  - odstraňují se bílé znaky na začátku a konci hodnoty atributu
  - sekvence bílých znaků se převádí na jednu mezislovní mezeru.

# Rozdíly oproti HTML 4 (pokračování)

- Skripty a elementy se specifikací stylu
  - musí být označena jako CDATA sekce:

```
<script type="text/javascript">
  <![CDATA[
    ... vlastní kód v javascriptu ...
  ]]>
</script>
```

- SGML vyloučení
  - SGML narozdíl od XML umožňuje v DTD zakázat určitý druh obsahu elementu
  - např. vnořené odkazy (elementy <a>)
- Elementy s atributy id a name
  - HTML 4 umožňovalo u některých elementů použít jako identifikátory fragmentů jak atribut name, tak id (a, applet, form, ...)
  - XHTML dokumenty musí používat atribut id.

# Rozdíly oproti HTML 4 (pokračování)

- Atributy s předdefinovanými hodnotami
  - XHTML je case sensitive, takže např. hodnoty TEXT a text jsou různé (element input).
- Odkazování entit pomocí hexadecimálních hodnot:
  - HTML umožňovalo oba zápisy &#xnn; i &#Xnn; v XHTML musí být &#xnn;
- Podrobnosti viz specifikace.

# Otázky kompatibility (1)

- Může být vhodné udržovat zpětnou kompatibilitu s HTML 4
  - Ne všechny prohlížeče korektně zpracovávají XHTML.
  - Je možné použít typ obsahu text/html místo application/xml+html, pokud se dodrží následující doporučení.
- Sada doporučení pro zpětnou kompatibilitu:
  - Pozor na to, že instrukce pro zpracování mohou být některými prohlížeči zobrazeny.:
  - Vkládejte mezeru před / a >
    - <br />, <hr />, 
  - Nepoužívejte minimalizovanou formu elementů (prázdné elementy):

- místo `<p/>` používejte `<p> ... </p>`
- ...

## Otázky kompatibility (2)

- Vyvarujte se konců řádků v hodnotách atributů.
- Používejte oba atributy `lang` a `xml:lang` pro specifikaci jazyku elementu.
- Pozor na identifikátory fragmentů.
- Specifikace kódování
  - používejte také element `meta`
- Boolovské atributy
  - některé prohlížeče nejsou schopny zpracovat plnou formu boolovských atributů.
- Další viz specifikace.

## Další zdroje informací

- [Specifikace HTML 4.01](#)
- [Specifikace XHTML 1.0](#)
- [Specifikace XHTML 1.1](#)
  - Vznikl modularizací XHTML 1.0 (moduly pro formuláře, hodnoty atributy elementu `li`, prezentaci XHTML, elementy a atributy pro styly, ..., více viz [specifikace](#))
- [Tutoriál na w3schools](#)
- XHTML [Strict](#), [Transitional](#) and [Frameset](#) Reference na [zvon.org](http://zvon.org).
- Řada knih:
  - S. M. Schafer: HTML, XHTML a CSS: bible
  - E. Castro: HTML, XHTML a CSS: Názorný průvodce tvorbou WWW stránek
  - P. Druska: CSS a XHTML: tvorba dokonalých webových stránek krok za krokem.