

Úvod do XML

(zdroj: <http://w3schools.com>)

Čo to je XML

- XML - **EX**tensible **M**arkup **L**anguage
- XML je značkovací jazyk rovnako ako HTML
- XML bol navrhnutý pre prenášanie a ukladanie dát, nie pre ich zobrazenie
- XML **značky nie sú preddefinované**. Používateľ si vytvára vlastné značky.
- XML je navrhnutý ako samoopisný jazyk.
- XML je W3C Recommendation (odporúčaním) od 10.2.1998

Rozdiely medzi XML a HTML

XML nie je náhrada za HTML.

XML a HTML boli navrhnuté s rozdielnymi cieľmi:

- XML bol navrhnutý pre **prenášanie** (transport) **a ukladanie dát** so zameraním sa na samotné dáta, čiže čo tieto dáta sú, čo obsahujú
- HTML bol navrhnutý pre **zobrazenie dát** so zameraním na zobrazenie týchto dát, čiže ako tieto dáta vyzerajú zobrazené

HTML sa zaoberá zobrazovaním informácií, XML sa zaoberá ich prenášaním a ukladaním.

XML kód nevykoná nič,

pretože XML bol vytvorený pre **štruktúrovanie, ukladanie a transport informácií**.

Nasledujúci príklad je správa od Petra pre Janu uložená v XML kóde:

```
<sprava>
  <odosielatel>Peter</odosielatel>
  <prijemca>Jana</prijemca>
  <titulok>Pripomienka</titulok>
  <telo>Nezabudni prist pred divadlo v sobotu.</telo>
</sprava>
```

Táto správa je úplne **samoopisná**. Má odosielateľa a prijímateľa správy (informácie), má tiež titulok a telo správy.

Ale tento samotný **XML kód nevykoná nič**. Je to len **informácia zabalená v značkách**. Musí existovať softvér, ktorý pošle, prijme alebo zobrazí túto informáciu (správu).

Ako vidieť z tohto príkladu XML kódu, XML dovoľuje autorovi XML dokumentu definovať jeho **vlastné značky** a jeho **vlastnú štruktúru** tohto **dokumentu**.

Naproti tomu značky v HTML dokumente sú preddefinované a môžeme v ňom používať **LEN** značky definované štandardom HTML.

Definícia XML (podľa W3C)

XML je softvérovo a hardvérovo nezávislý nástroj pre **prenášanie informácií**.

Inak povedané, XML je najobecnejší nástroj pre prenos dát medzi všetkými druhmi aplikácií.

Ako môže byť XML použité?

XML separuje dáta zobrazované pomocou HTML kódu **od samotného HTML kódu** do **separátnych XML súborov**, takže vývojár sa môže sústrediť na použitie HTML len pre zobrazenie a rozloženie dát v zobrazení. Môže sa spoľahnúť, že zmeny v separátne uložených dátach v XML súboroch nebudú vyžadovať zmeny v HTML kóde. Takto môže pomocou JavaScript kódu prečítať dáta z externého XML súboru a aktualizovať dátový obsah jeho webovej stránky.

XML zjednodušuje zdieľanie dát

Reálne počítačové systémy a databázy obsahujú dáta v nekompatibilných formátoch. **XML dáta** sú uložené v **čistom textovom formáte**, v **ASCII** súboroch. To poskytuje softvérovo a hardvérovo nezávislý spôsob ukladania dát a tiež to uľahčuje ich vytváranie. XML dáta môžu byť zdieľané rozdielnymi aplikáciami.

Vymieňanie dát vo formáte XML výrazne redukuje zložitosti, problémy pri výmene dát medzi rôznymi vzájomne nekompatibilnými aplikáciami cez Internet.

XML je použitý na vytváranie nových internetových jazykov

Veľa nových internetových jazykov je vytvorených pomocou alebo v XML:

- XHTML
- WSDL pre popis dostupných webových služieb for describing available web services
- WAP a WML sú značkovacie jazyky pre ručné zariadenia (handheld devices)
- RSS jazyky pre dodávanie nových správ
- RDF a OWL pre popisovanie zdrojov a ontológie
- SMIL pre popisovanie multimédií pre web

XML strom

XML dokumenty majú formu stromovej štruktúry, ktorá začína "koreňom" (root) a vetví sa do "listov" (leaves).

Príklad XML dokumentu

XML dokumenty používajú samoopisnú a jednoduchú syntax:

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<sprava>
  <odosielatel>Peter</odosielatel>
  <prijemca>Jana</prijemca>
  <titulok>Pripomienka</titulok>
  <telo>V sobotu nezabudni prist pred divadlo.</telo>
</sprava>
```

Prvý riadok je XML deklarácia, ktorá definuje XML verziu (1.0) a použité znakové kódovanie (ISO-8859-1 = Latin-1/West European character set).

Ďalší riadok popisuje **koreňový element** dokumentu (akoby hovoriaci „tento dokument je spáva“):

<sprava>

Ďalšie 4 riadky popisujú **4 detské elementy** koreňového elementu (odosielateľ, prijemca, titulok, telo):

```
<odosielateľ>Peter</odosielateľ>
<prijemca>Jana</prijemca>
<titulok>Pripomienka</titulok>
<telo>V sobotu nezabudni prísť pred divadlo.</telo>
```

Posledný riadok definuje koniec koreňového elementu:

```
</sprava>
```

Z tohto jednoduchého príkladu môžeme usudzovať, že tento XML dokument obsahuje správu od Petra pre Janu (samoopisnosť XML dokumentu).

XML dokumenty majú stromovú štruktúru

XML dokumenty **musia obsahovať koreňový element**. Tento element je „rodičom“ všetkých ostatných elementov.

Elementy v XML dokumente tvoria strom dokumentu. Strom začína koreňom a vetví sa do najnižšej úrovne stromu.

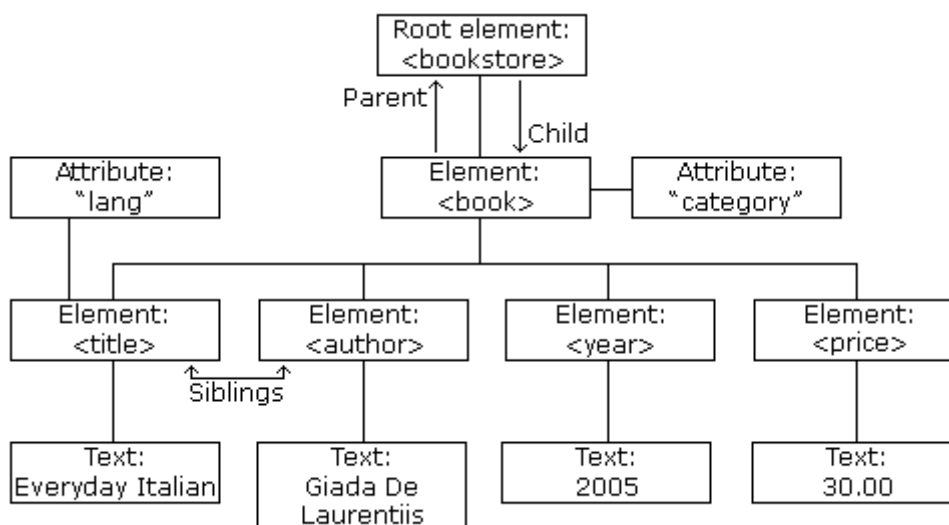
Všetky elementy môžu mať podelementy (detské elementy):

```
<root>
  <child>
    <subchild>.....</subchild>
  </child>
</root>
```

Pojmy rodič, dieťa a súrodenec sú použité na popisovanie vzťahov medzi elementmi. Rodičovské elementy majú deti. Deti na rovnakej úrovni sú nazvané súrodencami (bratia alebo sestry).

Všetky elementy môžu mať textový obsah a atribúty (rovnako ako v HTML).

Príklad vzťahov medzi elementmi v XML dokumente:



Obrázok vyššie reprezentuje jednu knihu (prvú) v nasledujúcom XML dokumente:

```
<bookstore>
  <book category="COOKING">
    <title lang="en">Everyday Italian</title>
    <author>Giada De Laurentiis</author>
    <year>2005</year>
    <price>30.00</price>
  </book>
  <book category="CHILDREN">
    <title lang="en">Harry Potter</title>
    <author>J K. Rowling</author>
    <year>2005</year>
    <price>29.99</price>
  </book>
  <book category="WEB">
    <title lang="en">Learning XML</title>
    <author>Erik T. Ray</author>
    <year>2003</year>
    <price>39.95</price>
  </book>
</bookstore>
```

Koreňovým elementom v tomto príklade je <bookstore>. Všetky <book> elementy v dokumente sú detskými elementmi koreňového elementu <bookstore>.

Element <book> má 4 detské elementy (deti): <title>, <author>, <year>, <price>.

Syntaktické pravidlá XML

Syntaktické pravidlá XML sú veľmi jednoduché a logické.

Všetky XML elementy musia mať uzatváraciu značku

```
<p>Toto je odstavec.</p>
<p>Toto je dalsi odstavec.</p>
```

XML elementy sú citlivé na výšku písma

```
<Sprava>Toto je nespravne</sprava>
<sprava>Toto je spravne</sprava>
```

XML elementy musia byť správne vnorené

```
<b><i>Tento text je tucnym a kurzivou</i></b>
```

Správne vnorenie elementov v tomto príklade znamená, že ak je element <i> otvorený vo vnútri elementu , tak musí byť aj uzavretý vo vnútri tohto elementu .

XML dokumenty musia mať koreňový element

XML dokumenty **musia obsahovať jeden element**, ktorý je rodičom všetkých ostatných elementov. Tento element sa nazýva **koreňovým elementom**.

```
<root>
  <child>
    <subchild>.....</subchild>
  </child>
</root>
```

Hodnoty XML atribútov musia byť v úvodzovkách

XML elementy môžu mať atribúty v názve alebo hodnote, ktoré však musia byť vždy v úvodzovkách.

Nesprávne použitie atribútu:

```
<sprava datum=13/2/2013>
  <odosielatel>Peter</odosielatel>
  <prijemca>Jana</prijemca>
  <titulok>Pripomienka</titulok>
  <telo> V sobotu nezabudni prist pred divadlo.</telo>
</sprava>
```

Správne použitie atribútu:

```
<sprava datum="13/2/2013">
  <odosielatel>Peter</odosielatel>
  <prijemca>Jana</prijemca>
  <titulok>Pripomienka</titulok>
  <telo> V sobotu nezabudni prist pred divadlo.</telo>
</sprava>
```

Niektoré znaky majú v XML špeciálny význam

Toto bude generovať XML chybu:

```
<sprava>ak je mzda < 1000 Eur, potom</sprava>
```

Pre vytvorenie bezchybného XML dokumentu musíme znak "<" nahradiť referenciou na túto entitu:

```
<sprava>ak je mzda &lt; 1000 Eur, potom</sprava>
```

5 preddefinovaných entitných referencií:

<	<	menší ako
>	>	väčší ako
&	&	ampersand
'	'	apostrof
"	"	znak úvodzovky

Komentáre v XML

<!--Toto je komentár -->

XML elementy

XML dokument obsahuje XML elementy.

Čo je a čo obsahuje XML element?

XML element je všetko medzi jeho začiatočnou a koncovou značkou, vrátane týchto značiek.

XML element môže obsahovať:

- iné elementy,
- text,
- atribúty,
- alebo mix tohto uvedeného ...

```
<bookstore>
  <book category="CHILDREN">
    <title>Harry Potter</title>
    <author>J K. Rowling</author>
    <year>2005</year>
    <price>29.99</price>
  </book>
  <book category="WEB">
    <title>Learning XML</title>
    <author>Erik T. Ray</author>
    <year>2003</year>
    <price>39.95</price>
  </book>
</bookstore>
```

V uvedenom príklade XML dokumentu elementy <bookstore> a <book> **obsahujú iné elementy**. Elementy <book> majú **atribúty** (napr. category="CHILDREN"). Elementy <title>, <author>, <year> a <price> obsahujú **text**.

Pravidlá pre vytváranie názvov v XML

XML elementy musia dodržiavať nasledujúce pravidlá pre vytváranie ich názvov:

- názvy môžu obsahovať písmená, čísla a iné znaky
- názvy nemôžu začínať číslom alebo interpunkčným znakom
- názvy nemôžu začínať písmenami xml (alebo XML, alebo Xml, atď.)
- názvy nemôžu obsahovať medzery

Hocijaké názvy môžu byť použité, žiadne slová nie sú rezervované.

XML atribúty

XML elementy môžu mať atribúty, ktoré poskytujú prídavné informácie o elemente.

XML validácia

XML dokument s korektnou syntaxou je správne vytvorený ("Well Formed" XML).

XML dokument validovaný oproti DTD (Document Type Definition) je validný (platný) ("Valid" XML).

Správne vytvorené XML ("Well Formed" XML) dokumenty

majú korektnú XML syntax.

Syntaktické pravidlá boli popísané vyššie:

- XML dokumenty musia mať koreňový element
- XML elementy musia mať uzatváracie značky
- XML značky sú citlivé na veľkosť písma
- XML elementy musia byť správne vnorené
- hodnoty XML atribútov musia byť v úvodzovkách

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<sprava datum="13/2/2013">
  <odosielatel>Peter</odosielatel>
  <prijemca>Jana</prijemca>
  <titulok>Pripomienka</titulok>
  <telo> V sobotu nezabudni prist pred divadlo.</telo>
</sprava>
```

XML Validator

testuje syntax XML dokumentu, dodržanie syntaktických pravidiel pri jeho vytváraní.

XML Errors Will Stop You

Syntaktické chyby v XML dokumente zastavia XML aplikáciu.

Prehliadanie XML súborov

je možné vo väčšine webových prehliadačov, avšak XML súbory nebudú nimi zobrazené ako HTML stránky.

Prehliadanie XML súboru

```
<?xml version="1.0" encoding="ISO-8859-1"?>
- <sprava datum="13/2/2013">
  <odosielatel>Peter</odosielatel>
  <prijemca>Jana</prijemca>
  <titulok>Pripomienka</titulok>
  <telo>V sobotu nezabudni prist pred divadlo.</telo>
</sprava>
```

Prečo je XML súbor takto zobrazený?

Pretože XML dokument neobsahuje informácie o tom, ako majú byť jeho dáta zobrazené.

Pretože XML značky sú vytvorené autorom XML dokumentu, prehliadač ich nepozná a teda ich ani nevie zobrazíť.

Bez akýchkoľvek informácií ako zobrazíť dáta XML dokumentu, väčšina prehliadačov zobrazí tento dokument tak, ako je.

Zobrazenie XML dokumentu pomocou CSS

Pomocou CSS (Cascading Style Sheets) mi môžeme pridať zobrazovacie informácie do XML dokumentu. Čiže je možné použiť CSS na formátovanie XML dokumentu.

Zobrazenie XML dokumentu pomocou XSLT

Pomocou XSLT môžeme transformovať XML dokument do HTML. XSLT je W3C odporúčaným štýlovým jazykom XML dokumentov. XSLT (eXtensible Stylesheet Language Transformations) je sofistikovanejší ako CSS. XSLT môže byť použitý na transformáciu XML dokumentu do HTML predtým, ako bude takto transformovaný dokument zobrazený webovým prehliadačom. Čiže XSLT transformácia je prehliadačom urobená vtedy, keď prehliadač číta XML dokument.
