

# Značkovací jazyky

Značkovací jazyk (anglicky markup language) slouží k obohacení textu o dodatečné informace – nejčastěji o významu, struktuře a způsobu zobrazování jednotlivých částí textu. Dodatečné informace se vkládají přímo do textu v podobě značek (tags), příkazů (commands) nebo direktiv. Výsledný zdrojový text je stále obyčejný textový soubor, což umožňuje jeho editaci i nejjednoduššími textovými editory. Nejznámějšími značkovacími jazyky jsou HTML (HyperText Markup Language) a XML, ve kterých je vytvořena většina WWW stránek.

## Původ

Pojem značkování (anglicky markup) souvisí se středověkou praxí připisování poznámek do rukopisů a s korektorskými značkami, sloužícími k opravám a úpravám textu před tiskem. Autor dával tiskařovi značkou na vědomí, že bude změna typu písmen, že má vynechat místo na obrázek atd.

## Příklad zápisu

Pro vymezení řídicích konstrukcí – značek či příkazů – slouží znaky se speciálním významem. Například v HTML a XML to jsou znaky „menší než“ (<) a „větší než“ (>), jež ohraničují značky. Text mezi nimi je chápán jako instrukce, kterou využije zpracovávající software. Zdrojový text může vypadat například takto:

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

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

Výsledný text by se zobrazil takto:

## Nadpis

Text prvního odstavce se *zvýrazněným* slovem.

## Příklady značkovacích jazyků

### TeX

Mezi hlavní publikační standardy patří systém TeX vytvořený ve druhé polovině 70. let 20. století Donaldem Knuthem se záměrem vytvořit nástroj pro naprosto přesný popis rozvržení textu pro sazbu odborné literatury v profesionální kvalitě. V průběhu 80. let byl doplněn mimo jiné o systém pro popis znaků Metafont a systém pro souběžný vývoj programového vybavení a dokumentace (Web). TeX je díky svým rozšířením používán pro sazbu odborné i krásné literatury a je standardním publikačním formátem v mnoha odvětvích vědy. K zápisu značek používá lomítka a složené závorky.

```
\chapter{Začátek kapitoly}
```

Text uvnitř kapitoly.

Výsledný text by se zobrazil takto:

## Začátek kapitoly

Text uvnitř kapitoly.

### GML

GML (Generalized Markup Language) je vůbec první obecný značkovací jazyk, který se trochu prosadil. Byl publikován v roce 1973 a v roce 1974 byl na jeho základě vytvořen jazyk SGML.

### SGML

SGML (Standard Generalized Markup Language) byl jeden z nejrozšířenějších značkovacích jazyků. Měl ohromné možnosti ve vytváření dalších formátů. Kládl (v dnešní době zbytečný) důraz na možnost úspory paměti tím, že obsahoval řadu výjimek a nepravidelností, což způsobilo, že je velmi těžké napsat program zpracovávající text v takovémto formátu. V praxi se stejně všechny jeho možnosti nepoužívaly. V dnešní době je nahrazen jazykem XML.

```
<DOKUMENT>
  <adresa stát="cs">
    <město>Brno</město>
    <ulice>
  </adresa>
</DOKUMENT>
```

### XML

XML (Extensible Markup Language) je zjednodušená verze jazyka SGML. Vznikl odstraněním jeho

nepotřebných prvků a jeho celkovou modernizací. Definice typu dokumentu se zde může vytvořit pomocí mnoha k tomu určených formátů. Jednou z největších výhod aplikací XML je jejich vzájemná snadná převeditelnost. Dnes je běžné vytvořit dokument např. v DocBooku (jedna z aplikací XML) a pak jej převést na spoustu různých formátů pro různá média. Jako jediný z obecných značkovacích jazyků se používá i samostatně na internetu (tvůrce stránky si vlastně vytvoří nový datový formát podle svých potřeb). Se samotným XML dokumentem však autor musí publikovat i jeho DTD soubor a soubor definující vzhled prvků (pomocí jazyků CSS nebo XSL FO). Toto řešení má ohromné možnosti, jeho problémem však je, že jej podporují jen novější internetové prohlížeče, naučit se ho používat je také mnohem složitější, než staré jednoduché HTML. Samozřejmě se na internetu objevuje i několik všeobecně uznávaných formátů, které vznikly jako aplikace XML: např. XHTML, RSS, atd.

```
<?xml version="1.0" encoding="UTF-8" ?>
<recept jméno="chleba_s_máslem_se_salámem" příprava="3 minuty">
  <titulek>Chleba s máslem a se salámem</titulek>
  <přísada množství="1" jednotka="krajíce">Krajíc chleba</přísada>
  <přísada množství="1" jednotka="kousek">Máslo</přísada>
  <přísada množství="2" jednotka="plátky">Salám</přísada>
</recept>
```

## HTML

HTML (HyperText Markup Language) je asi nejznámější aplikací SGML. Je to základní nástroj pro vytváření internetových stránek. Jeho výhodou je, že se dá dost snadno naučit a podporují jej úplně všechny internetové prohlížeče. Text je doplněn značkami (tagy), popisujícími buď jeho význam nebo vzhled. Postupem času byl jazyk HTML doplněn kaskádovými styly (CSS), které umožňují definovat vzhled dokumentu ještě mnohem přesněji a lépe. 28. října 2014 byla vydána finální specifikace verze HTML5, která přidala například několik nových značek nebo podporu přehrávání multimédií v prohlížeči bez nutnosti použít Flash.

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

```
<p>Text prvního odstavce s <strong>tučným</strong> slovem.</p>
```

Výsledný text by se zobrazil takto:

# Nadpis

Text prvního odstavce s **tučným** slovem.

## XHTML

Extensible HyperText Markup Language (XHTML) vznikl úpravou jazyka HTML. Je striktnější, což se dá považovat za jeho výhodu oproti klasickému HTML. Například vyžaduje používání malých písmen u tagů, nepovoluje klasický zápis nepárových značek (nepárové značky musí být ukončeny mezerou a lomítkem – například `<br>` se v XHTML musí zapsat jako `<br />`), vyžaduje hodnotu u všech atributů tagu (například `<input type=„checkbox“ checked>` se v XHTML musí zapsat jako `<input type=„checkbox“ checked=„checked“ />`) a všechny hodnoty atributů musí být zapsány v uvozovkách. Původně měl nahradit HTML, ale vzhledem k tomu, že poslední finální verze je XHTML 1.1 z roku 2001, je zřejmé, že XHTML klasické HTML nenahradí.

## RSS

RSS (Rich Site Summary) je moderní formát (aplikace XML) využívaný zejména na serverech publikujících velké množství článků. Obsahuje jen stručnou charakteristiku nových článků a odkazy na ně. Je zpracováván speciálním programem (některé nové prohlížeče jej mají zabudovaný). Formát RSS je využíván zejména pro svou přehlednost a rychlost.

## SVG

SVG (Scalable Vector Graphics) je aplikace XML pro kreslení vektorovou grafikou.

## Využití

Značkovací jazyky sice byly původně určeny pro textové dokumenty, ale stále více se používají i v jiných oblastech pro prezentaci různých typů informací, mimo jiné seznamů skladeb, vektorové grafiky, webových služeb, content syndication a uživatelských rozhraní. Většinou se jedná o XML aplikace, protože XML je dobře definovaný a rozšiřitelný jazyk.

## Verze HTML

### HTML do verze 3.2

Původní HTML (Tim Berners-Lee, 1990). Pouze základní tagy, většinou dodnes funkční.

Ve výsledku si autoři webových prohlížečů HTML sami pozměnili. Postupně se přidaly formuláře, tabulky, barvy, pozadí, později také podpora Javascriptu, nepočítalo se však s CSS.

### HTML 4.01

Nové tagy (`<div>`, `<span>`) a zjednodušení díky plnohodnotnému zavedení CSS (a tedy i oddělení smyslu a vzhledu stránky).

Reálně se (špatně) úprava vzhledu v HTML a v CSS často kombinuje. V rámci správného zápisu je rovněž dobré dbát na rozdíl mezi vizuálními a logickými tagy. Například: vizuální tag `<b>` textu říká,

že má být tučný, kdežto jeho logická obdoba `<strong>` říká, že je daný text důležitý, a proto je v základu tučně zvýrazněn.

HTML 4.01 mělo nahradit XHTML, které bylo striktnější, ale neuchytilo se. Rozdíly ve striktnosti HTML 4.01 vystihují i jeho 3 typy DTD (Document Type Definition, obsahuje informace o použitém druhu HTML): **Strict** (neobsahuje prezentační tagy typu `<font>`), **Transitional** (obsahuje prezentační tagy), **Frameset** (obsahuje prezentační tagy a navíc přidává framesety - velmi nevhodné tagy pro rozdělení stránky na několik dílčích rámců).

## HTML 5

Vývoj po zklamání ve formě XHTML pokračoval primárně vylepšováním CSS (dnes verze 3). Až po roce 2010 se začalo objevovat HTML 5.

HTML 5 je založeno na striktní verzi HTML 4.01, přidává ale mnoho nových tagů. Například tagy pro přehlednější strukturování stránek (`<section>`, `<article>`, `<header>`, `<footer>`, `<nav>`) a pro multimédia (`<video>`, `<audio>`) a vektorovou grafiku (`<canvas>`).

Deklarace: `<!DOCTYPE html>`. (Podstatně kratší než u HTML 4.01, protože není nutné specifikovat DTD).

From:

<http://wiki.gml.cz/> - GMLWiki

Permanent link:

<http://wiki.gml.cz/informatika:maturita:14a?rev=1581079771>

Last update: **07. 02. 2020, 13.49**

