

ZNAČKOVACÍ JAZYKY A JEJICH VYUŽÍVÁNÍ

MARKUP LANGUAGE AND THEIR USE

Zdeněk Havlíček

Anotace:

Značkovací jazyk vkládá do textu dokumentu další informace o vlastním textu. Značkovací jazyky se využívají především pro přenos informací (dat) mezi počítači. Jsou vhodným nástrojem pro popis nejrůznějších dokumentů. Pomocí značek se popisuje jak interní struktura dokumentu, tak se uvádějí i externí charakteristiky datového souboru. Značkovací jazyk XML je základem pro nové technologie webových služeb, které se stále více využívají při inovaci podnikových informačních systémů.

Klíčová slova:

Značkovací jazyky, HTML, XML, XHTML, webové služby

Abstract:

A markup language inserts into the given text of a document, additional, descriptive information about the text itself. Markup languages are used primarily for the transfer of information (data) between computers. They are a useful tool for the description of a variety of documents. By using tags, the internal structure of a document is described and external characteristics of the data file are also shown. The XML markup language is the basis for the new technologies of web services, which are increasingly being used for innovation of company information systems.

Key words

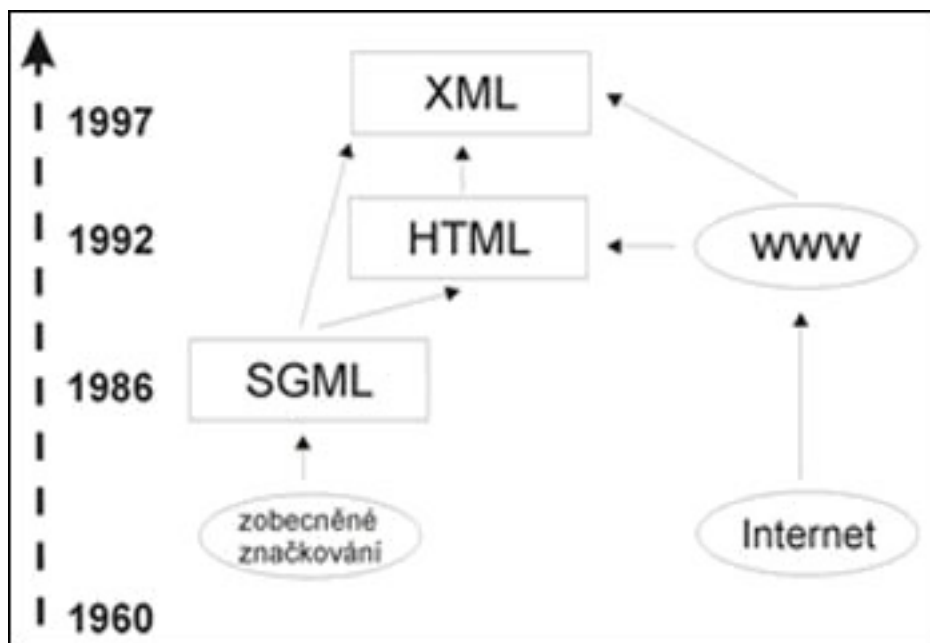
Markup languages, HTML, XML, XHTML, web services

ÚVOD

Značkovací jazyk je jakýkoli jazyk, který pomocí speciálních značek vysvětluje význam (sémantiku) různých částí textu nebo určuje vzhled (formát) jednotlivých částí textu. Značky se původně používaly jen pro úpravu (korekturu) dokumentu při závěrečném formátování knih v nakladatelstvích.

Koncepce sémantického značkování dat vnikla již koncem 60. let. Později, v roce 1986, byl organizací ISO schválen obecný základ pro značkovací jazyky, tzv. SGML (Standard Generalized Markup Language).

Nové informační technologie na začátku devadesátých let umožnily praktickou aplikaci značkovacího jazyka. Byl vytvořen jazyk HTML (Hypertext Markup Language), který se stal jedním ze tří stavebních pilířů www. V 1998 byla formálně uvedena první verze obecného značkovacího jazyka, nazvaného XML (eXtensible Markup Language). XML navazuje na principy a konvence jazyků SGML a HTML (viz obr. č.1).



Obr. č. 1 – Vývoj značkovacích jazyků

CÍL

Cílem příspěvku je především ukázat na výhody a nevýhody značkovacích jazyků a současně zhodnotit nejpoužívanější značkovací jazyky z hlediska dalšího vývoje podnikových informačních systémů.

VÝSLEDKY

Na základě studia literatury a vlastních praktických zkušeností je vhodné značkovací jazyky rozdělit na tři samostatné skupiny:

- Jazyk HTML
- Jazyk XML
- Jazyk XHTML.

HTML

Jazyk HTML (Hypertext Markup Language, hypertextový značkovací jazyk) je především nástroj na tvorbu a zobrazování hypertextových dokumentů ve WWW.

- Využívá se k vyznačení struktury informace v elektronickém dokumentu. Pomocí značek HTML lze téměř libovolně strukturovat vlastní text dokumentu.
- Definiuje syntaxi a rozmístění speciálních vložených příkazů (značek), které prohlížeč nezobrazuje, ale které řídí způsob zobrazení dokumentu.
- Popisuje způsob zobrazení dokumentu v prohlížeči příjemce informace, tzn., že určuje vzhled, a tím i význam jednotlivých prvků na stránce.

Standard jazyka HTML prošel od svého vzniku v roce 1991 překotným vývojem a jeho původní účel – zpřístupnění vědeckých informací a dokumentů na síti www – byl dalece překročen. Se zaváděním Internetu do dalších oblastí lidské činnosti se začalo měnit i využívání jazyka HTML.

Zvětšující se rozmanitost prostředí, kde je jazyk HTML využíván, a narůstající problémy s kompatibilitou podnítily potřebu usměrnit vývoj webových technologií směrem k

větší pružnosti a lepší rozšiřitelnosti. Vývoj nových verzí jazyka HTML od roku 1995 koordinuje sdružení *World Wide Web Consortium* (W3C). Současným standardem, který se ale dále již nevyvíjí, je verze HTML 4.01, která byla definována konsorciem W3C v roce 1999.

Zatímco různé značky a pravidla jazyka HTML jsou dobře definované, většina dodavatelů prohlížečů poskytuje pro tento jazyk rozšíření nad definice W3C. Prohlížeče samotné jsou příliš „volné“ při dodržování pravidel značkového jazyka, což vede k „nedbalému“ využívání této technologie.

XML

Jazyk XML vychází také z jazyka SGML. Zatímco jazyk HTML je zaměřen od svého vzniku především na definování způsobu rozvržení dat na stránce, v jazyce XML lze data popisovat typem informací, které reprezentují. XML není pouze dalším značkovacím jazykem, ale komplexním systémem pravidel pro definování, vytváření a sdílení dat.

XML je vlastně metajazyk, což znamená, že je to prostředek používaný k definování jiných jazyků. Značky musí být v XML organizovány podle určitých obecných principů, ale co se týká významu značek, dává nám XML naprostou volnost. Značky, které se v XML vytvoří, mohou být zdokumentovány v tzv. DTD (Document Type Definition) nebo v tzv. schématech.

Při používání XML dokumentu potřebujeme také dokument zobrazit. XML nemá žádné specifické prostředky pro definici vzhledu. Existuje ale několik stylových jazyků, které umožňují definovat, jak se mají jednotlivé elementy zobrazit. Jeden vytvořený styl můžeme aplikovat na mnoho dokumentů stejného typu, stejně tak můžeme na jeden dokument aplikovat několik různých stylů. Prezentace dokumentu (vzhled) se potom definuje připojeným stylem. Další možností je pomocí různých stylů provést transformaci do jiného typu dokumentu, nebo do jiné struktury XML. Výsledkem může být např. soubor typu pdf.

Jazyk XML je určen především pro výměnu dat mezi aplikacemi a pro publikování dokumentů. Jazyk umožňuje popsat strukturu dokumentu z hlediska věcného obsahu jednotlivých částí, nezabývá se sám o sobě vzhledem dokumentu nebo jeho částí. XML umožňuje vytvářet bohatší www stránky, ke kterým lze přistupovat z různých prohlížečích platforem (mobilní telefony, televizory, atd.).

Příklady aplikací XML:

- RDF - Resource Description Framework, specifikace, která umožňuje popsat externí vlastnosti souborů např. obsah a anotaci dokumentu;
- SMIL - Synchronized Multimedia Integration Language, popisuje multimediální soubory pomocí XML.
- MathML - Mathematical Markup Language je značkovací jazyk pro zápis matematických vzorců a symbolů;
- SVG - Scalable Vector Graphics je jazyk pro popis dvourozměrné vektorové grafiky,
- DocBook - Sada definic dokumentů a stylů pro publikační činnost.

XHTML

Konsorcium W3C vydalo v roce 2000 doporučení pro využití jazyka XML, označené jako XHTML (Extensible Hypertext Markup Language). Jazyk XHTML byl vytvořen s důrazem na dobrou rozšiřitelnost, kompatibilitu a přenositelnost obsahu popsaného tímto značkovacím jazykem.

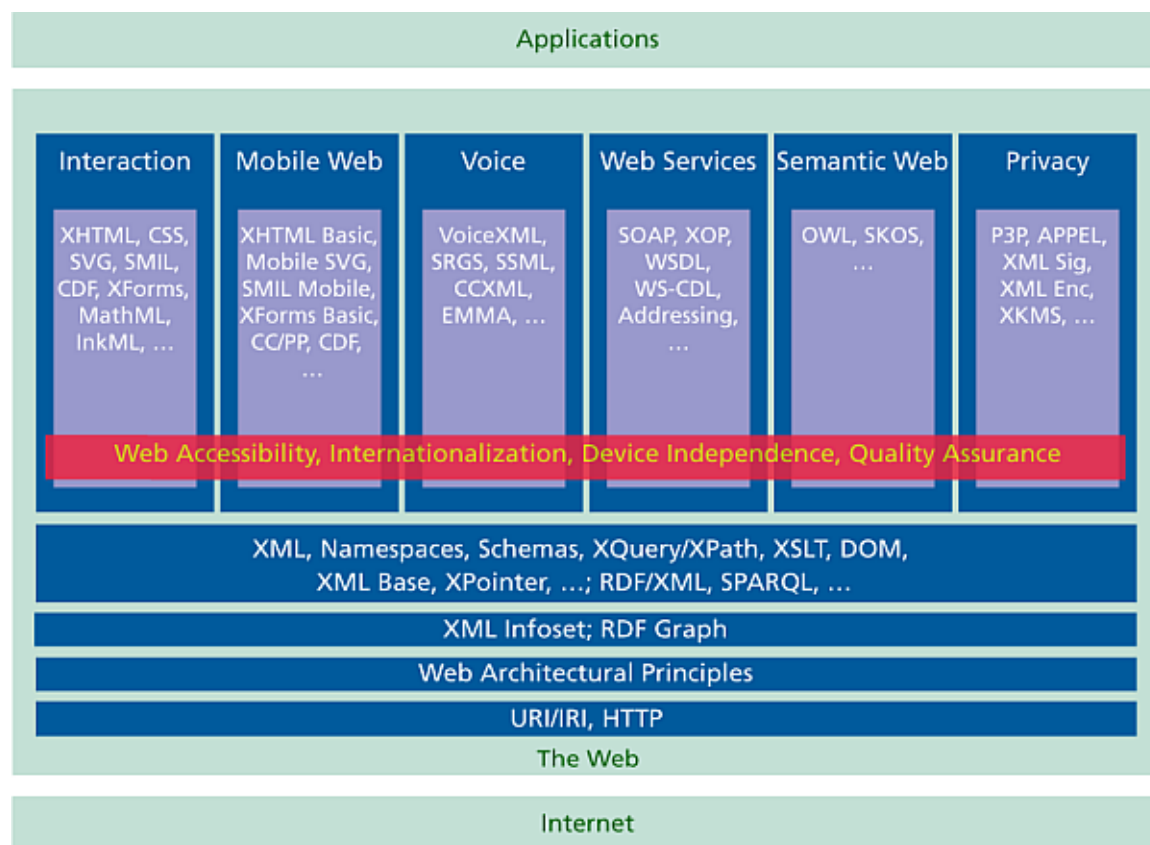
Jazyk XHTML představuje v podstatě přísnější a syntakticky čistší verzi HTML. Definice typu dokumentu `<!DOCTYPE ...>` je povinná. Dokumenty musí být well-formed (značky se zapisují pouze malými písmeny, jednotlivé prvky musí být korektně vnořovány, všechny párové značky musí být ukončeny, prázdné značky musí být ukončeny jako v XML, např. jako ``). Jazyk XHTML jasně odděluje vzhled dokumentu od jeho struktury, tím že podporuje využívání CSS (Cascading Style Sheets).

ZÁVĚRY

Hlavní vlastnosti značkovacích jazyků jsou tyto:

- Značkovací jazyky jsou nezávislé na platformě. Jsou dobře přenosné mezi různými operačními systémy a nejsou závislé ani na způsobu zobrazení (na tom mohou být závislé až jejich aplikace);
- Značky (anglicky tags) vkládané do textu jsou člověkem čitelné, což usnadňuje kontrolu správnosti;
- Značkovací jazyky obsahují jen informace o struktuře dokumentu (ne o jeho formátování). Díky tomu mohou být zobrazené v mnoha médiích (počítač, mobil, tiskárna, ...).

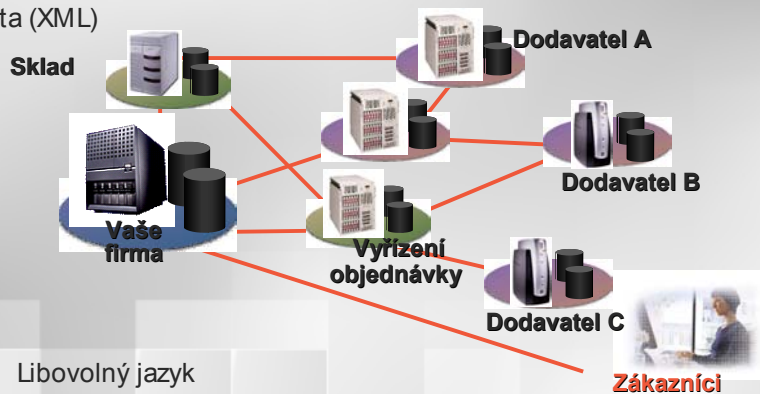
Značkovací jazyky a především jejich typický a nejnovější zástupce jazyk XML se stávají základem dalšího vývoje WWW (viz obr. č.2) a přenositelnosti dat mezi stroji pomocí webových služeb. Webové služby umožňují automatizovanou výměnu dat mezi různými informačními systémy a tak se stávají novou technologií, která bude stále častěji využívána (viz obr. č. 3).



Obr. č. 2 – Technologie www – dle konsorcia W3C

Web Services

- Najdi (UDDI)
- Popiš (WSDL)
- Zavolej (SOAP)
- Data (XML)



- Libovolný jazyk
- Libovolná platforma
- Libovolný dodavatel

Obr. č. 3 – Technologie pro webové služby

LITERATURA

Harold, E.R., Means, W.S. [2004]: XML in a Nutshell. O'Reilly.

Havlíček, Z. [2004]: Web technology and farmers. *Agricultural Economics* 50(6): p. 243-247. ISSN 0139-570X.

Sobotková, K. [2004]: Intranet – význam a příklad řešení. PEF ČZU.

Šilerová, E. [2004]: Informační systémy – nástroj změn. *Proceedings AP XIII*. CZU Praha. ISBN 80-213-1190-8.

WWW.w3.org

WWW.oasis.com

cs.wikipedia.org

Kontakt:

Doc. Ing. Zdeněk Havlíček, CSc.

vedoucí KIT PEF ČZU v Praze

e-mail: Havlicek@pef.czu.cz

tel: 224 382 273