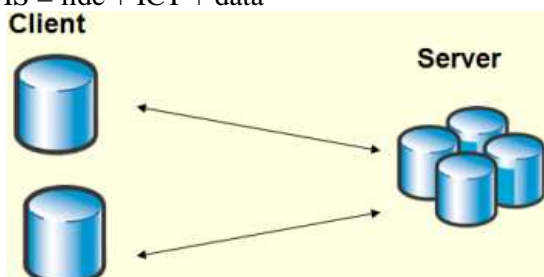


INTERNETOVÉ TECHNOLOGIE – CLIENT SIDE P2

2008-10-07

ZNAČKOVACÍ JAZYKY:

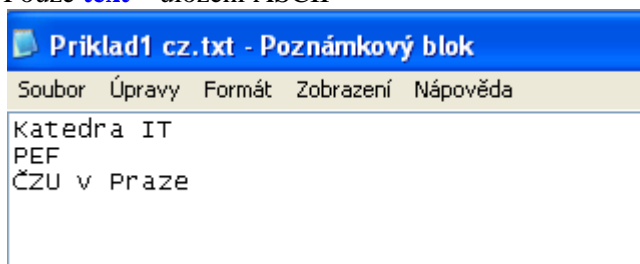
- ✓ IS = lidé + ICT + data



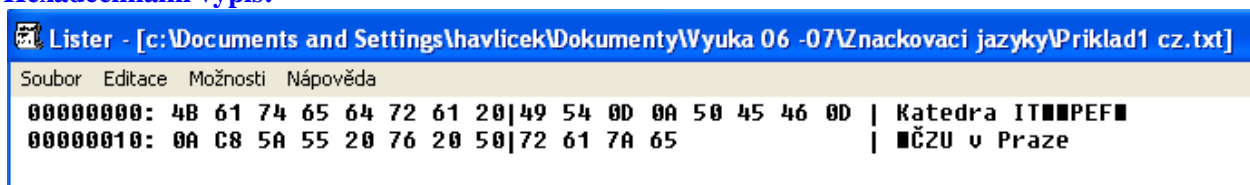
- ✓ Www – více než 50 000 000 000 statických stránek

Jak ukládat data:

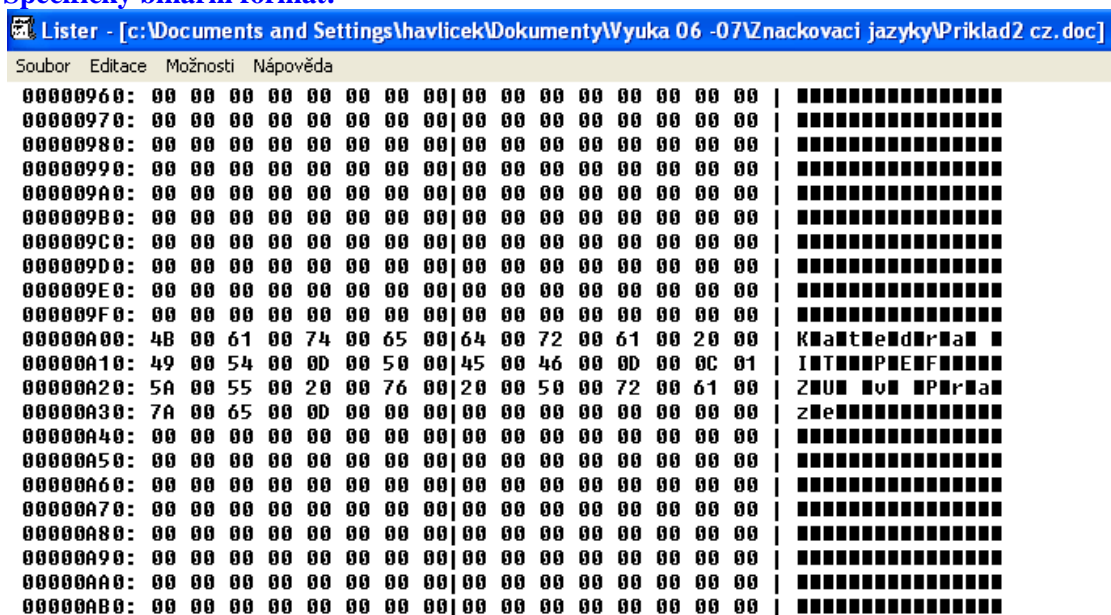
- ✓ Pouze **text** – uložení ASCII

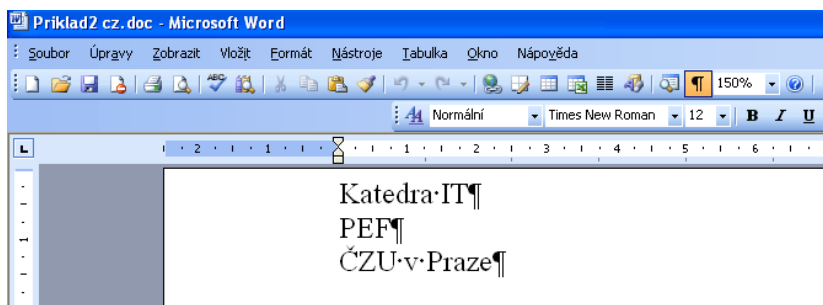


- ✓ **Hexadecimální výpis:**

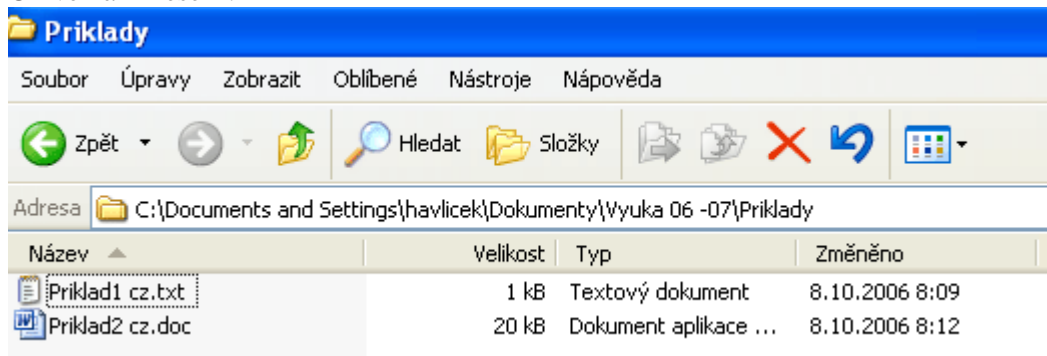


- ✓ **Specifický binární formát:**





- ✓ Univerzální řešení?



- ✓ Vývoj ICT – nárůst výpočetních kapacit
- ✓ Pouze text => značkovací jazyky

Charakteristika:

- ✓ Značkovací jazyk vkládá do textu dokumentu další **informace o vlastním textu**. Značkovací jazyk kombinuje **text a informaci** o textu.
- ✓ Značkovací jazyky se využívají především pro **přenos informací** (dat) mezi počítači.
- ✓ Jsou dobře přenosné mezi různými operačními systémy, nejsou na žádném OS závislé.
- ✓ **Značkovací jazyk ≠ programovací jazyk**. Ve značkovacím jazyku popisujeme data, ale nedáváme instrukci jako v programovacím jazyku.

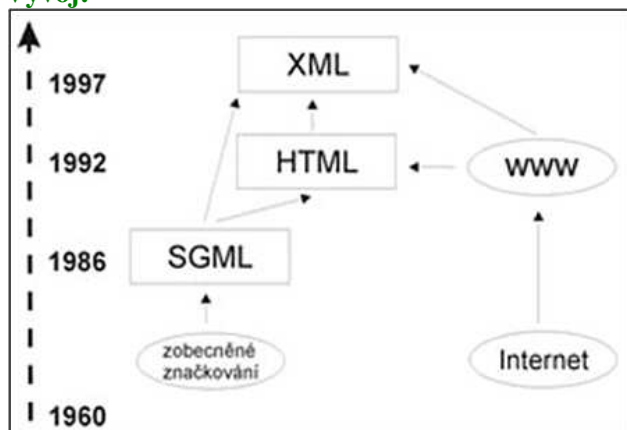
Historie:

- ✓ **Značky** (angl. Markup) se původně používaly jen pro úpravu (korekturu) dokumentu při závěrečném formátování knih v nakladatelstvích.
- ✓ **Korektorské** speciální **symbols** pro označení úprav v tištěném textu

Definice:

- ✓ Značkovací jazyk je jakýkoli jazyk, který pomocí speciálních značek vysvětluje **význam** (sémantiku) různých částí textu
- ✓ Značkovací jazyk může určovat **vzhled** (formát) jednotlivých částí textu.

Vývoj:





SGML:

- ✓ **SGML – Standard Generalized Markup Language**
- ✓ Obecný základ pro značkovací jazyky – norma ISO
- ✓ SGML is complicated to understand and difficult to integrate into an application.

HTML:

- ✓ Jazyk **HTML (Hypertext Markup Language)**, hypertextový značkovací jazyk) je značkovací jazyk určený pro tvorbu www stránek, které se zobrazují prostřednictvím prohlížeče.
- ✓ Využívá se především k **vyznačení struktury informace v elektronickém dokumentu**.
- ✓ HTML je obecný nástroj na tvorbu a zobrazování hypertextových dokumentů ve WWW.
- ✓ 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.
- ✓ Vzniklo několik verzí jazyka, poslední je 4.01
- ✓ Konsorcium W3C rozhodlo, že vzhledem k příchodu XHTML, vývoj HTML je ukončen – nevytváří se již žádná nová verze HTML.
- ✓ HTML is the **lingua franca** for publishing hypertext on the World Wide Web. It is a non-proprietary format based upon SGML, and can be created and processed by a wide range of tools, from simple plain text editors to sophisticated WYSIWYG authoring tools.
- ✓ **Lingua franca** je jakýkoliv jazyk šířeji využívaný nad rámec rodilých mluvčích.

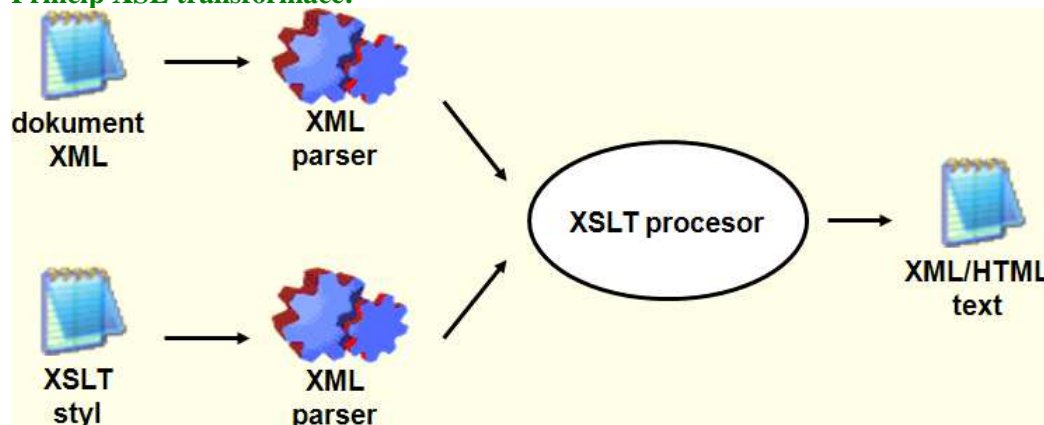
XML:

- ✓ **XML (eXtensible Markup Language)**, česky rozšiřitelný značkovací jazyk) je obecný značkovací jazyk, který byl vyvinut a standardizován konsorciem W3C.
- ✓ XML je metajazyk, což znamená, že je to prostředek používaný k definování jiných jazyků.
- ✓ XML není pouze dalším značkovacím jazykem, ale komplexním systémem pravidel pro definování a **sdílení dat**.
- ✓ **XML** umožňuje snadné vytváření konkrétních značkovacích jazyků pro různé účely a široké spektrum různých typů dat.
- ✓ **Jazyk 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.
- ✓ **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.

XSL:

- ✓ Jazyk **XSL (eXtensible Stylesheet Language)** má dvě funkční stránky:
 - **XSLT** (XSL Transformations) - slouží k transformaci dokumentů, umožňuje vytvářet styly.
 - **XSLFO** (XSL Format Objects) - část, která slouží k přesnému popisu vzhledu dokumentu.

Princip XSL transformace:



- ✓ **Parser** = syntaktický analyzátor



Výhody stylů:

- ✓ Možnost formátování jednoho dokumentu pomocí více různých stylů



XML – pojmy:

- ✓ **Značka (tag)** – identifikátor pro data, <klient></klient>
- ✓ **Element** – datová struktura včetně značek, <klient>Jan Novák</klient>
- ✓ **Atribut** – upřesňuje vlastnosti elementu, <klient id="123"></klient>
- ✓ **Instrukce** pro zpracování, <? xml:>
- ✓ **Comments** <!--comment-->

Data a metadata:

- ✓ V jazyce XML lze data popisovat typem informací, které reprezentují.
- ✓ Pomocí nových značek vytváříme tzv. metadata, tj. data o datech.
- ✓ Vytváříme data, která dávají význam jiným datům
- ✓ **Příklad – řešení:**

- <basnik>Pierre de Ronsard</basnik>
 <FrenchPoet>Pierre de Ronsard</FrenchPoet>
- <ruze>Pierre de Ronsard</ruze>
 <RoseCultivar>Pierre de Ronsard</RoseCultivar>

Příklad XML – stromová struktura:

```
<?xml version="1.0" encoding="windows-1250" ?>
- <kartaOdrudy>
  <jmenoOdrudy>Westerland</jmenoOdrudy>
  <typRuze>Floribunda. Sadová – parková</typRuze>
- <puvodOdrudy>
  <zeme>D</zeme>
  <slechtitel>Kordes 1969</slechtitel>
</puvodOdrudy>
- <kvet>
  <velikost>8-12 cm</velikost>
  <barva>základní barva oranžová s odstíny meruňkové a lososově růžové</barva>
  <vune>silná ****</vune>
</kvet>
- <rostlina>
  <vyska>do150cm</vyska>
  <typ>vzpřímená, košatá</typ>
  <odolnostChoroby>dobrá</odolnostChoroby>
</rostlina>
  <pouziti>Soliterní nebo v malých skupinách</pouziti>
  <oceneni>ADR 1974</oceneni>
</kartaOdrudy>
```



Příklad – Word 2003:

```

<?xml version="1.0" encoding="UTF-8" ?>
<kartaOdrudy >
  <jmenoOdrudy>Westerland</jmenoOdrudy>
  <typRuže>Floribunda·Sadová—parková</typRuže>
  <puvodOdrudy >
    <zeme>D</zeme>
    <slechtitel>Kordes·1969</slechtitel>
  </puvodOdrudy>
  <kvet >
    <velikost>8-12·cm</velikost>
    <barva>základní·barva·oranžová·s·odstíny·meruňkové·a·lososově·růžové</barva>
    <vune>silná·****</vune>
  </kvet>
  <rostlina >
    <vyska>do150cm</vyska>
    <typ>vzpřímená,·košatá</typ>
    <odolnostChoroby>dobrá</odolnostChoroby>
  </rostlina>
  <pouziti>Soliterní·nebo·v·malých·skupinách</pouziti>
  <oceneni>ADR·1974</oceneni>
</kartaOdrudy>

```



```

<?xml version="1.0" encoding="UTF-8" ?>
<kartaOdrudy >
  <jmenoOdrudy>Westerland</jmenoOdrudy>
  <typRuže>Floribunda·Sadová—parková</typRuže>
  <puvodOdrudy >
    <zeme>D</zeme>
    <slechtitel>Kordes·1969</slechtitel>
  </puvodOdrudy>
  <kvet >
    <velikost>8-12·cm</velikost>
    <barva>základní·barva·oranžová·s·odstíny·meruňkové·a·lososově·růžové</barva>
    <vune>silná·****</vune>
  </kvet>
  <rostlina >
    <vyska>do150cm</vyska>
    <typ>vzpřímená,·košatá</typ>
    <odolnostChoroby>dobrá</odolnostChoroby>
  </rostlina>
  <pouziti>Soliterní·nebo·v·malých·skupinách</pouziti>
  <oceneni>ADR·1974</oceneni>
</kartaOdrudy>

```

Příklad:

Westerland

Typ růže	-	Floribunda. Sadová – parková
Původ	země	D
	šlechtitel	Kordes 1969
Květ	velikost	8-12 cm
	barva	základní barva oranžová s odstíny meruňkové a lososově růžové
	vůně	silná ****
Rostlina	výška	do 150 cm
	typ	vzpřímená, košatá
	odolnost proti chorobám	dobrá
Použití	-	Soliterní nebo v malých skupinách
Ocenění	-	ADR 1974
Foto		
		

XML – shrnutí:

- ✓ XML představuje obecný standard pro značkování dokumentů
- ✓ Je to metajazyk (a markup language) pro textové dokumenty



- ✓ Data v XML dokumentech jsou ve formě textových řetězců
- ✓ Základní jednotka při značkování se označuje jako element
- ✓ XML umožňuje se používat pro:
 - **Narrative-Like Documents** – pro záznamy jednání, více „upovídané“
 - **Record-Like Documents** – jako záznamy, pro databázové technologie
- ✓ DTD – Document Type Definition

Technologie XML:

- ✓ **Narrative-Like Documents:**
 - XHTML
 - TEI (Text Encoding Initiative) – klasické texty
 - DocBook – manuály, nové knihy
 - OpenOffice
 - ...
- ✓ **Record-Like Documents:**
 - Databases
 - Web services - SOAP
 - XML Schema
 - RSS
- ✓ **Pro práci s XML dokumenty:**
 - XSLT
 - XML schema
 - DOM
 - RSS

Silné stránky XML:

- ✓ Jednoduchá syntaxe
- ✓ Hierarchická struktura elementů
- ✓ Snadná kontrola správnosti- značky jsou čitelné pro člověka – human/readable
- ✓ Jazyk nezávislý na platformě – využívá kodování Unicode

Slabé stránky XML:

- ✓ Rychlý vývoj
- ✓ Příliš velké množství aplikací
- ✓ Na dílčích standardech se podílí řada vývojových skupin – např. OASIS

XHTML:

- ✓ **Extensible HyperText Markup Language**, nebo **XHTML**. Značkovací jazyk má stejné charakteristiky jako HTML
- ✓ XHTML má přísnější pravidla pro zápis značek
- ✓ XHTML je aplikací XML

Zápis značek:

- ✓ Dokumenty musí být „well-formed“, tj. strukturovány podle pravidel
- ✓ Značky jsou párové – vytváří element (objekt)
`<mama> </mama>`
`<a> `
- ✓ Značky se zapisují pouze malými písmeny
- ✓ Jednotlivé prvky musí být korektně vnořovány
- ✓ Všechny párové značky musí být ukončeny
- ✓ Prázdné značky musí být ukončeny

Závěr:

- ✓ Značkovací jazyky jsou nezávislé na platformě
- ✓ Značkovací jazyky využívají značky (human-readable tags)



- ✓ Technologie využívající značkovací jazyky se rozšiřují do nerůznějších oblastí (vektorová grafika, web services, RSS).

