



DIPLOMOVÁ PRÁCA

Zuzana Valentová

**Tvorba a publikácia matematickej
terminológie pomocou XML**

Bratislava 2006

**FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY
UNIVERZITY KOMENSKÉHO V BRATISLAVE
Mlynská dolina, 842 48 Bratislava**

KATEDRA APLIKOVANEJ INFORMATIKY



DIPLOMOVÁ PRÁCA

Tvorba a publikácia matematickej
terminológie pomocou XML

Bratislava 2006

Diplomant: Zuzana Valentová
Konzultant: Doc. RNDr. Andrej Ferko, PhD.

Zadanie diplomovej práce

Študent(ka)..... *Zuzana Valentová*.....

vypracuje za účelom obhajoby v rámci záverečného konania vo vysokoškolskom magisterskom štúdiu na UK FMFI diplomovú prácu s (predbežným) názvom: *Tvorba a publikácia matematickej terminológie pomocou XML*

pod vedením pracovníka..... *Doc. Andrej Ferko* z pracoviska..... *KAGDM*

Cieľ diplomovej práce a ďalšie poznámky:

Spracovať rukopis matematickej terminológie, navrhnúť formu publikácie, implementovať a otestovať v matematickej komunite.

podpis študenta..... podpis ved. dipl. práce.....

Diplomová práca je tematicky priradená k špecializácii: *Počítačová grafika*.....

V Bratislave dňa..... *30. 09. 2004* podpis gestora špec.:.....

Poznámka: Vyplňujte v 3 exemplároch: 1 pre študenta, 1 pre vedúceho práce, 1 pre študijné oddelenie (odovzdá študent pri zápise do 4. ročníka).

Vyhlasujem, že som túto prácu vypracovala samostatne, len s použitím uvedenej literatúry.

V Bratislave 15.05.2006

*Život má ten zmysel,
ktorý mu dáme.*

Za pomoc a podporu ďakujem
Doc. RNDr. A. Ferkovi, PhD.,
mamine,
Ľubici,
Zdenke
a Martinovi.

Abstrakt

Táto práca sa zaoberá otázkou ako zobrazíť matematickú terminológiu na internete alebo na CD. Zameriava sa na riešenie s použitím jazyka *XML* a jazyka *MathML*. Uvádza návod ako postupovať pri spracovaní vstupných materiálov a publikuje výsledné dielo online. Prácu podporil grant KEGA č. 3/0010/02 Dokončenie matematickej terminológie Terminologickou komisiou JSMF, ktorá poskytla 660 strán materiálov.

Obsah

Úvod	10
1 Extensible Markup Language	
1.1 História <i>XML</i>	12
1.2 Základné pojmy	13
1.3 <i>XML</i> verzus <i>HTML</i>	15
1.4 Základné prvky <i>XML</i> dokumentu	15
1.5 <i>MathML</i>	17
2 Terminológia	
2.1 Význam terminológie	19
2.2 Základné pojmy z oblasti terminológie	20
2.3 Matematická terminológia	21
3 XML Editor	
3.1 Špecifikácia	24
3.2 Popis funkcií	25
4 Postup práce	
4.1 Postup	30
4.2 <i>XML</i> súbory	34
5 Internetová stránka	
5.1 Štruktúra	39
5.2 Obsah	40
5.3 Analýza výsledkov ankety	45

6 Záver	50
7 Literatúra	52
8 Príloha A	
Definície pojmov a skratiek	55
9 Príloha B	
Obrazová príloha	58
10 Príloha C	
DVD – obsah	60

Zoznam obrázkov, tabuliek a príkladov

Príklad 1.1:	Jednoduchý <i>XML</i> dokument.	15
Tabuľka 1.2:	Základné matematické značky.	17
Príklad 1.3:	Zápis vzorca pomocou <i>XML</i> a <i>MathML</i> .	18
Obrázok 3.1:	Hlavné okno <i>XML Editora</i> .	24
Príklad 3.2:	Formát vstupného textového súboru.	26
Príklad 3.3:	Výstupný <i>XML</i> súbor.	26
Obrázok 3.4:	Tlačidlá na vkladanie matematických značiek.	27
Obrázok 3.5:	Tlačidlá na vkladanie gréckej abecedy.	27
Obrázok 3.6:	Tlačidlá na vkladanie špeciálnych znakov.	28
Obrázok 3.7:	Tlačidlá na vkladanie šípok.	28
Obrázok 3.8:	Tlačidlá na vkladanie matematických operátorov.	29
Obrázok 4.1:	Postup práce.	30
Príklad 4.2:	Formát textového súboru.	31
Príklad 4.3:	<i>XML</i> súbor transformovaný z textového súboru.	32
Príklad 4.4:	<i>XML</i> súbor po úprave.	33
Obrázok 4.5:	Zobrazenie <i>XML</i> súboru na internete.	34
Príklad 4.6:	Štruktúra <i>XML</i> súborov.	35
Príklad 4.7:	Štruktúra súboru <code>pocitacova_grafika.xml</code> .	36
Tabuľka 5.1:	Zoznam <i>XML</i> súborov s počtom termínov.	40
Obrázok 5.2:	Formulár na vkladanie termínov do databázy.	41
Tabuľka 5.3:	Štruktúra tabuliek na ukladanie termínov z formulára.	42
Obrázok 5.4:	Formulár na vyhľadávanie termínov.	42
Tabuľka 5.5:	Štruktúra tabuľky <code>terminy</code> .	43
Tabuľka 5.6:	Štruktúra tabuľky <code>anketa</code> .	44
Tabuľka 5.7:	Štruktúra tabuľky <code>kniha</code> .	44
Tabuľka 5.8:	Štatistický prehľad odpovedí z ankety.	49

Úvod

Každé nové riešenie rodí nové problémy.

MURPHYHO ZÁKON

Každá krajina, každý štát, každá národnosť sa odlišuje svojim jazykom. Taktiež aj história vzniku jazyka a písma je odlišná. Slovenčina je prirodzený a živý jazyk, reč nádherná. No na to, aby dosiahla túto dnešnú podobu bolo potrebných mnoho rokov a mnoho ľudí, ktorým záležalo na jej napredovaní a vývoji. V dnešnej dobe, keď veda a technika učinili obrovské pokroky, musí slovenský jazyk taktiež neustále reagovať na ich zrýchlený vývoj.

Našou snahou bolo prispieť v oblasti matematickej terminológie. Matematika má široké uplatnenie vo všetkých sférach spoločenského života, osobitne v značne matematizovaných blízkyh prírodných vedách a v technických odboroch. Len spoľahlivo reprezentované matematické znalosti umožňujú inžinierom, technikom, konštruktérom i majstrom umenia neustále držať krok s technickým rozvojom a s matematickou presnosťou plniť požiadavky na ne kladené. Predpokladom k tomu je solídne matematické vzdelanie.

Cieľom našej práce bolo vytvoriť komplexný systém matematickej terminológie a terminológie z počítačovej grafiky pomocou jazyka XML. Navrhnuť elektronickú formu publikácie, implementovať a publikovať výsledné dielo na internete, aby bolo prístupné nielen širokej matematickej komunite ale aj všetkým používateľom internetu a výsledné dielo otestovať v matematickej komunite. Na rozdiel od stredoškolských učebníc [16] sme sa snažili obohatiť našu prácu aj o interaktivitu, t. j. o vyhľadávanie konkrétnych termínov. Práca vznikla s podporou grantu KEGA č. 3/0010/02 Dokončenie matematickej terminológie Terminologickou komisiou JSMF, ktorá nám poskytla materiály [14].

V prvej kapitole sme zhrnuli základné pojmy jazykov *XML* a *MathML*. Zamerali sme sa najmä na nesporné výhody použitia týchto jazykov. V druhej kapitole sme podali stručný prehľad o tvorbe terminológie. V tretej kapitole sme opísali program, ktorý sme vytvorili a použili v procese vytvárania terminológie a ktorý je nevyhnutnou súčasťou postupu tvorby. V ďalšej kapitole sme sa zamerali na opis konkrétnych krokov postupu práce a štruktúry výsledných *XML* súborov. V poslednej kapitole sme opísali výsledné dielo, ktoré sa nachádza na internetovej stránke [25] a zhodnotili anketu, ktorá analyzuje odozvu v matematickej komunite.

1 Extensible Markup Language

S pokrokom by nič nebolo, len nie a nie ho skončiť.

OGDEN–NASHOV ZÁKON

Kapitolu 1 sme spracovali z [18], [19] a [20].

1.1 História jazyka XML

Na začiatku 80–tych rokov minulého storočia bola snaha veľkých firiem dohodnúť na istom štandarde pre výmenu údajov medzi rôznymi počítačmi a tak vznikol štandard *ISO 8879* z roku 1986. Táto definícia obsahuje definíciu metajazyka *SGML (Standard Generalized Markup Language)*. Je to jazyk, ktorý umožňuje zdieľanie informácií medzi podnikmi s rozdielnymi informačnými systémami, resp. umožňuje oddeliť dáta od ich spracovania. *SGML* dokument je textový dokument, prenositeľný na rôzne systémy, ktoré majú implementovaný analyzátor *SGML* dokumentov. Analyzátor určí štruktúru a spracuje jeho obsah. Možnosti *SGML* boli využívané už pred nástupom internetu, vznikol v dobe, keď boli počítače pomalé a drahé. Po nástupe internetu sa ukázala jeho neschopnosť flexibilne reagovať na požiadavky webu. Ďalšími jeho nevýhodami sú komplexnosť, všeobecnosť a samotná implementácia je veľmi náročná. Časom sa ukázalo, že v praxi sa aj tak používa len určitá časť. Tato najdôležitejšia podmnožina bola vybraná ako nový jazyk a dostala meno *XML (eXtensible Markup Language)*.

Prvú verziu zverejnili vo februári 1998. V októbri roku 2000 vyšla revízia tohto odporúčania pod názvom *XML 1.0 (Second Edition)*. Akronym *XML* je skratkou anglického výrazu *eXtensible Markup Language*, ktorý by sme mohli do slovenčiny preložiť ako „rozšírený značkovací jazyk“. Ďalšou známou aplikáciou jazyka *SGML* je *HTML (Hyper Text Markup Language)*. Tento jazyk je v dnešnej dobe veľmi populárny hlavne kvôli jeho jednoduchosti.

1.2 Základné pojmy XML

Čo je to markup language? XML je tzv. markup language, značkovací jazyk určený pre dokumenty obsahujúce štruktúrované informácie. Pojem markup language znamená, že jazyk obsahuje nejakým spôsobom zavedený mechanizmus pre identifikáciu nielen samotných hodnôt dokumentu, ale aj ich význam. To znamená, že spolu obsahuje obsah + význam štruktúrovaných dát.

Čo je to metajazyk? XML patrí do rodiny tzv. metajazykov. Pojem meta všeobecne znamená povýšenie o jednu úroveň abstrakcie. Ďalej vyjadruje všeobecný vzťah medzi možnou abstrakciou a implementáciou tejto abstrakcie. Napríklad metatrieda v OOP je „trieda tried“, metamodel je „nad model“, kde jednou z jeho implementácií je daný model. Pretože jazyk je možné chápať ako súhrn pravidiel, potom metajazyk je nejakým konkrétnym súhrnom pravidiel pre tvorbu jazyka, t. j. jazyk jazyka. XML ako metajazyk deklaruje syntax pomocou tzv. značiek (tagov). Umožňuje vytvoriť ľubovoľnú značku a tým dáva možnosť vytvárať ľubovoľné štruktúry, a práve preto je chápaný ako metajazyk.

Čo je to dokument, štruktúra dokumentu? Jazyk XML sa používa na riadenie predovšetkým štruktúry dokumentu. Dokument je chápaný vo všeobecnosti. Niektorí používatelia si pojem dokument spájajú výhradne len s oblasťou publikovania. Toto terminologické nedorozumenie potom môže viesť k presvedčeniu, že XML je len novým riešením publikovania stránok pre WWW. Korene XML pochádzajú samozrejme z oblasti publikovania, ktorej už jazyk HTML nedáva dostatočné možnosti. Pričom opäť je vhodnejší širší pohľad. Publikujú sa nielen webové stránky, ale aj technická dokumentácia, listy, knihy, cenníky, faktúry, obchodné dokumenty... Pre tieto účely sa XML výborne hodí. Veľa veľkých i malých firiem zaoberajúcich sa výrobou softvéru, hardvéru alebo aj lietadiel používa pre tvorbu dokumentácie systémy založené na XML alebo SGML. Dokumenty vo formáte XML sú uložené ako textové dokumenty a dajú sa priamo čítať a meniť ľubovoľným textovým editorom. Pri manipulácii s nimi sa často používajú kaskádové štýly (CSS). Dokumenty môžu navzájom vstupovať do rôznych vzťahov, ktoré

sa dajú vyjadriť pomocou odkazov. Kľúčom k pochopeniu *XML* je myšlienka, že dokumenty majú štruktúru, obsah a tiež veľmi často aj formát. Tieto tri zložky dokopy vytvárajú dokument. Takisto sa navzájom ovplyvňujú a pri práci s dokumentom sa dajú jednoducho zameniť.

Štruktúra je akýsi plán dokumentu, zachytáva jeho logickú organizáciu. Definuje, z čoho sa dokument skladá, v akom poradí sa majú prvky vyskytovať a sú určené aj ich vzťahy. Uvažujeme o klasických textových dokumentoch, textové procesory prácu so štruktúrou v uvedenom zmysle nepodporujú a ponechávajú všetko na autorovi. V *XML* ide o to, aby bola štruktúra explicitne definovaná. Autor rozhoduje o tom, či bude štruktúra veľmi striktná alebo naopak voľnejšia.

Obsah je samotné naplnenie dokumentu údajmi. Dokumenty sa skladajú z textu, jednotlivých slov, prípadne ilustrácií, t. j. ikonických a syntaktických reprezentácií významu.

Formát špecifikuje, ako sa jednotlivé slová, vety, odstavce vizuálne prezentujú a navzájom rozlišujú v rámci dokumentu. Špecifické formátovanie slúži k rôznym účelom – zvýraznenie, vizualizácia štruktúry a celková estetická a typografická úprava dokumentu. Práve formát je často nesprávne zamieňaný so štruktúrou. Je to dané tým, že v textových procesoroch je štruktúra implikovaná formátovaním pomocou štýlov. Pomenovanie štýlu by malo zodpovedať postavenie daného bloku textu v rámci dokumentu. Niektorí používatelia niekedy zabúdajú a používajú rovnaký štýl pre formátovanie rôznych typov textov. Ešte horšia situácia nastáva, keď používateľ použije vo svojom dokumente veľa rôznych druhov písma, rôzne veľkosti písma a rôzne iné formátovacie charakteristiky. Výsledok je síce vizuálne zaujímavý, skôr však v negatívnom zmysle, ale veľmi obtiažne čitateľný a zrozumiteľný.

XML ponúka mechanizmus pre vymieňanie štruktúrovaných dokumentov, v ktorých je vždy zachovaná štruktúra spoločne s obsahom. Je to možné vďaka tomu, že definície štruktúry i obsahu sú oddelené od špecifického formátovania. Dokumenty, ktoré sú vyjadrené pomocou *XML* je možné preniesť do nového systému s minimálnym úsilím.

1.3 XML verzus HTML

XML súbor je uložený ako textový súbor, čítať sa dá ľubovoľným textovým editorom. V *XML* sa význam dát označuje pomocou zátvoriek `<` a `>`. Napríklad pri knihe môžeme zaviesť značku *ISBN* takto:

```
<isbn>80-251-0085-5</isbn>
```

Po sémantickej stránke je prístup pomocou značiek podobný ako v *HTML*. Oproti *HTML* je tu však dôležitý rozdiel. Zatiaľ čo *HTML* je jazyk, *XML* je *metajazyk*. To znamená, že *HTML* má obmedzenú množinu značiek, t. j. syntax tohto jazyka je daná jeho značkami a ich možnosťami. Pokiaľ zavedieme v *HTML* značku, ktorá nie je zadaná, tak bude ignorovaná alebo spôsobí chybu. A keďže *HTML* je len „jazyk“, ktorému musí rozumieť napr. internetový prehliadač, tak každá definícia novej značky musí byť premietnutá aj do funkcionality prehliadača. *XML* nemá množinu značiek obmedzenú a navyše nie je dopredu stanovený význam týchto značiek. Napr. značka `` v *HTML* slúži na vloženie obrázku, v *XML* môže mať ľubovoľný význam, napr. imaginárna časť čísla.

1.4 Základné prvky XML dokumentu

```
1. <?xml version = "1.0"?>
2.     <kniha jazyk = "sk">
3.         <titul>C# Programujeme profesionálne</titul>
4.         <autor>
5.             <meno>Simon</meno>
6.             <priezvisko>Robinson</priezvisko>
7.         </autor>
8.         <nakludaje>
9.             <nakladatel>Wrox</nakladatel><br />
10.            <rok>2003</rok>
11.         </nakludaje>
12.         <isbn>80-251-0085-5</isbn>
13.     </kniha>
14. </xml>
```

Príklad 1.1:

Jednoduchý *XML* dokument.

Inštrukcie na spracovanie slúžia na pridanie informácií na spracovanie dokumentu, povahovo však nie sú obsahom dokumentu. Uzatvárajú sa medzi zátvorky `<? a ?>` (Príkl. 1.1, riadok 1). Inštrukcie sa najčastejšie používajú na pridanie štýlu, ktorý definuje vzhľad, na zaradenie príkazov pre rôzne preprocesory a pod.

Elementy sa nachádzajú v každom XML dokumente, môžu byť navzájom vnorené. Elementy sú vyznačené párové značky (začiatočná a koncová) a zapisujú sa medzi znaky `< a >` (Príkl. 1.1, riadok 2), pričom koncová značka obsahuje pred názvom ešte znak `/` (Príkl. 1.1, riadok 13). Niektoré elementy nemusia mať žiadny obsah, napr. element pre zalamovanie riadkov stačí zapísať značkou `
` namiesto `
</br>` (Príkl. 1.1, riadok 9). Elementy musia byť korektne uzatvorené a musia byť úplne vnorené do iného elementu.

Atribúty definujú dodatočné vlastnosti elementu a zapisujú sa vždy v počiatočnej značke elementu (Príkl. 1.1, riadok 2). Jeden element môže mať viacero atribútov.

Komentáre môžu byť súčasťou dokumentu, zapisujú sa do zátvoriek `<!-- a -->`. Požívajú sa na vloženie akejkoľvek poznámky. Tiež môžu slúžiť na skrytie niektorej časti textu. Pri spracovávaní dokumentu sú ignorované.

Celý XML dokument má textový obsah. Skladá sa z vlastných znakových dát a zo značiek, oboje sú textové reťazce. Tým sa zaručuje zrozumiteľnosť formátu. Na zápis znakov sa používa znaková sada ISO [31], ktorá je 32 – bitová, čo znamená, že môže obsahovať až 2^{32} znakov. Je možné používať aj iné kódovania. Dokumenty s iným kódovaním musia vždy začínať deklaráciou XML so zodpovedajúcim atribútom, napr.:

```
<?xml version = "1.0" encoding = "UTF-8"?>
```

Entita je chápaná ako pomenovaná časť textu. Entitou môže byť celý dokument, ľubovoľná sekvencia znakov alebo ľubovoľný symbol, ktorý sa nachádza v dokumente. Entita je zahrnutá do dokumentu pomocou odkazu, ktorý sa začína znakom `&` a končí `;`.

1.5 MathML

Matematika vyjadruje myšlienky aj pomocou symbolov, ktoré nie sú jednoducho zobraziteľné napr. v internetových prehliadačoch. *MathML* je značkovací jazyk, ktorý je určený práve na zápis matematických symbolov. *MathML* je podmnožinou *XML*, preto pri vkladaní matematických symbolov stačí použiť textový editor. Štruktúra obsahuje tie isté prvky ako *XML* a to elementy, atribúty, entity. Tabuľka 1.2 obsahuje niektoré zo základných značiek.

<mtext>	normálny text
<mspace />	medzera
<mi>	identifikátor premennej
<mn>	čísla
<mo>	operátory
<math>	matematický výraz
<table>	tabuľka
<mrow>	horizontálna skupina prvkov
<mtd>	bunka tabuľky
<mfrac>	zlomok
<msqrt>	druhá odmocnina
<mroot>	všeobecná odmocnina
<msub>	dolný index
<sup>	horný index
<msubsup>	horný aj dolný index
<munder>	obsah pod daným prvkom
<mover>	obsah nad daným prvkom
<munderover>	obsah nad aj pod daným prvkom

Tabuľka 1.2:
Základné matematické značky.

Príklad 1.3 znázorňuje zápis nasledujúceho výrazu pomocou *MathML*:

$$\frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

```
<?xml version = "1.0" encoding = "iso-8859-1"?>
  <math xmlns = "http://www.w3.org/1998/Math/MathML">
    <mfrac linethickness="1">
      <mrow>
        <mo>-</mo>
        <mi>b</mi>
        <mo>&plusmn;</mo>
        <msqrt>
          <msup>
            <mi>b</mi>
            <mn>2</mn>
          </msup>
          <mo>-</mo>
          <mn>4</mn>
          <mi>a</mi>
          <mi>c</mi>
        </msqrt>
      </mrow>
      <mrow>
        <mn>2</mn>
        <mi>a</mi>
      </mrow>
    </mfrac>
  </math>
```

Príklad 1.3:

Zápis vzorca pomocou *XML* a *MathML*.

2 Terminológia

Omnogo dôležitejšie ako fakty sú názvy, ktoré im dáš.
COHENOV ZÁKON

Podkapitolu 2.1 a 2.2 sme spracovali z [9], [10] a [11].

2.1 Význam terminológie

Terminológiu môžeme chápať ako disciplínu zameranú na teoretické problémy tvorby pojmov v rámci určitej tematickej oblasti, reprezentácie pojmových systémov a problémy odborných jazykov. Alebo ako štruktúrovaný súbor pojmov a ich označení v určitej tematickej oblasti.

Terminológia (*Terminology science*) je interdisciplinárna tematická oblasť zameraná na pojmy, ktoré sa považujú za základné jednotky myslenia (tvorba pojmov) a poznania. Ponúka metódy na tvorbu, zaznamenanie, spracovanie a viacnásobného využitie terminologických údajov. Ako vedná disciplína využíva poznatky filozofie, vedy o vede, aplikovanej lingvistiky, komunikačných a informačných vied a aj ďalších aplikovaných vied, ako sú teórie klasifikácie, informácie a dokumentácia a normalizácia. Zaoberá sa štúdiom:

- pojmov a vzťahov medzi nimi
- pojmových systémov
- reprezentácie pojmov prostredníctvom definícií a označení
- modelov tvorby termínov
- frazeologických aspektov odborných jazykov
- princípov terminologickej činnosti

Terminológia je súčasťou vývinu každého vedného odboru, ovplyvňuje teóriu aj prax. Je potrebné usmerňovať jej vytváranie a predísť tak živelnému vývinu. Terminologická činnosť je vo svete čoraz viac uznávaná. Nárast potreby údajov a terminologických metód vyžaduje rozvoj normalizácie v oblastiach zameraných na tvorbu terminologických databáz.

2.2 Základné pojmy z oblasti terminológie

Terminologická práca je zameraná na systematické usporiadanie a sprístupnenie poznatkov. Základným prvkom je pojem, pojmy sa analyzujú a štrukturalizujú. Súčasťou terminologickej práce je:

- excerpčia
- harmonizácia
- terminografia

Excerpcia je činnosť, pri ktorej sa vyberajú terminologické údaje prehľadávaním súborov textov alebo iných jazykových údajov zhromaždených na analýzu.

Harmonizácia pojmov je činnosť zameraná na redukovanie menej podstatných rozdielov medzi dvoma alebo viacerými pojmi, ktoré spolu súvisia. Harmonizácia pojmov je súčasťou štandardizácie. Harmonizácia termínov je činnosť zameraná na označenie jedného pojmu v rôznych jazykoch termínmi, ktorých význam je rovnakej alebo podobnej charakteristiky.

Terminografia je činnosť zameraná na zaznamenanie a prezentáciu terminologických údajov. Terminológia môže byť prezentovaná v podobe terminologických bánk, terminologických slovníkov, prekladových slovníkov a pod. Terminologická banka je banka dát obsahujúca terminologické údaje a terminologická databáza je typ bázy dát, ktorá obsahuje terminologické údaje.

Označenie (pomenovanie) je reprezentácia pojmu lingvistickými alebo nelingvistickými prostriedkami. V terminologickej práci označeniami môžu byť termíny, názvy alebo symboly, ktoré pojem označujú, vzťahujú sa na pojem, alebo sú pojmu priradené. Typy označení pojmu:

verbálne

- alfanumerické
- grafické
- obrazové

neverbálne

- skratka
- termín (všeobecný pojem)
- názov (jednotlivý pojem)

Termín je verbálne označenie všeobecného pojmu v špecifickej tematickej oblasti a *názov* je verbálne označenie jednotlivého pojmu. *Symbol* je označenie pojmu prostredníctvom písmen, piktogramov alebo ich kombinácií.

Terminologický manažment sa vo všeobecnosti chápe ako cieľavedomá manipulácia s terminologickými údajmi. Systém terminologického manažmentu môžeme chápať ako softvérový nástroj na spracovanie terminologických údajov. Z hľadiska počítačového spracovania terminológie je dôležitý formát terminologického záznamu. Výsledkom terminologickej práce je terminologický produkt v tradičnej tlačenej forme alebo v elektronickej forme.

Súčasný rozvoj informačných technológií umožňuje integráciu rôznych typov údajov, medzinárodnú výmenu terminologických informácií a viacnásobné využitie informácií. Vytvárajú sa tak predpoklady a aj potreby na koordináciu terminologických aktivít na medzinárodnej úrovni. Nutným predpokladom viacnásobného využitia terminologických informácií je zjednotenie metodológie na vytváranie terminologických údajov.

Podkapitolu 2.3 sme spracovali z [12] a [13].

2.3 Matematická terminológia

Vývoj jazyka, didaktickej teórie a terminologickej teórie neustále napreduje. Autor pri vytváraní matematickej terminológie musí neustále sledovať vývoj v oblasti jazyka a terminológie a neustále sa zdokonaľovať v ich ovládaní. Pri svojej práci môže využiť špecializované publikácie jazykových odborníkov o tvorbe a používaní terminológie. Nedostatočná znalosť základov teórie terminológie vedie k presvedčeniu, že tvorba termínov je voluntaristická činnosť, v ktorej zavedenie termínu spočíva v dohode istého okruhu ľudí. Je to však náročný proces, ktorý by mal zachovávať *princíp vedeckosti* a realizovať ho pomocou základnej zásady odbornej terminológie – *princípu jednoznačnosti termínov*.

Výrazové prostriedky matematickej terminológie pomenúvajú matematické objekty, relácie, operácie, atď. Tieto pomenovania tvoria jadro matematickej terminológie a podstatnú časť jej objemu. Sú slovným pomenovaním matematických pojmov, ktoré môžu byť písomne zachytené špecifickým jazykom v zmysle matematickej logiky jej symbolickými prostriedkami. Predovšetkým si všetci používatelia odbornej terminológie (učitelia, žiaci, autori textov, čitatelia) majú uvedomiť obsah slova v prirodzenom jazyku a obsah toho istého slova ako termínu pomenúvacieho jednoznačne definovaný pojem vednej disciplíny.

Problémy a chyby v používaní matematickej terminológie majú svoj pôvod v nevedení si, resp. nerešpektovaní rozdielov medzi významom určitých slov v prirodzenom jazyku a ich obsahom v sústave terminológie niektorej odbornej oblasti, resp. niekoľkých oblastí. Cieľom je odlišenie hovorovej úrovne prirodzeného jazyka od exaktnej terminologickej úrovne jazyka matematiky, tým sa pestuje a podporuje presnosť myslenia, jednoznačnosť formulácií a môže sa predchádzať nedorozumeniam. Prenos zo zásoby prirodzeného jazyka do terminologickej sústavy je logicky a didakticky náročný krok. Napríklad aj také „neterminologické“ slovo prirodzeného jazyka ako je *udalosť*, je odborným termínom v teórii pravdepodobnosti a v matematickej štatistike.

Téma je zjednocujúci faktor, ktorý sa sústreďuje na využitie poznatkov z rôznych predmetov na komplexné spracovanie problematiky, na jej posúdenie z pohľadu niekoľkých odborov. Téma napomáha k vytvoreniu uceleného obrazu o sledovaných objektoch a vedie k uvedomeniu si súvislostí a významových rozdielov v pojmoch z rôznych vedných disciplín. Napríklad termín *usporiadanie* sa vyskytuje v teórii aritmetických a algebraických štruktúr a aj v elementárnej geometrii. No treba si uvedomiť rozdiel v definíciách. V prvom prípade je pojem *usporiadanie* chápaný ako *binárna relácia*, v druhom prípade ako *ternárna relácia v množine všetkých bodov priamky*.

3 XML Editor

Pracujú stroje, ľudia rozmýšľajú.

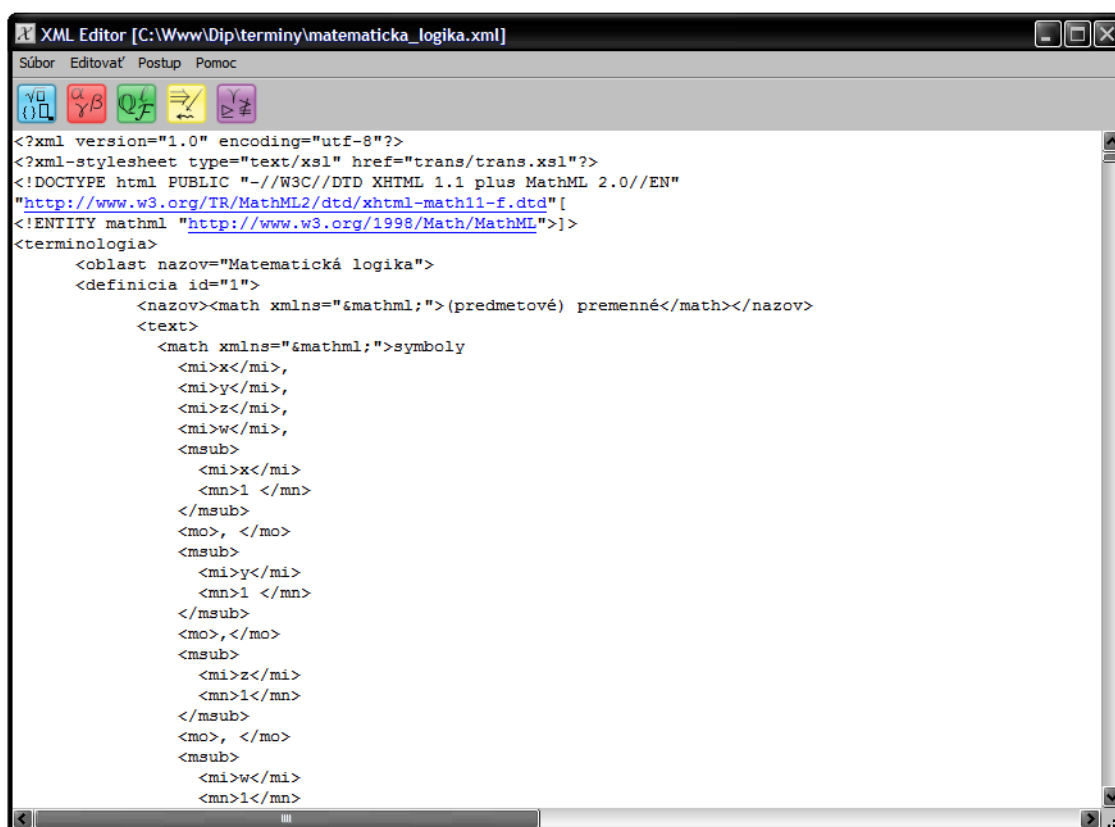
POLLYANOVA ZÁSADA

Pri spracovávaní matematickej terminológie sme vytvorili jednoduchý program, ktorý slúžil na transformáciu textového súboru do *XML* súboru. Neskôr sme dospeli k záveru, že nie je potrebné vytvorené textové súbory len do *XML* súboru transformovať, ale ich aj ďalej spracovávať. Manuálne vpisovanie matematických značiek je veľmi neefektívne, časovo náročné a je potrebné neustále kontrolovať korektnosť zápisu značiek. Preto sme sa rozhodli vytvoriť si vlastný editor, ktorý by vedel vkladať matematické značky.

Prečo vlastný? Existuje už niekoľko programov na prácu s *MathML* jazykom. Za najvhodnejší pokladáme program Amaya s otvoreným zdrojovým kódom vyvinutý konzorciom *W3C*. Tento nástroj v sebe zahŕňa *WYSIWYG* editor aj editor zdrojového kódu a prehliadač internetových stránok. Tento program podporuje štandardy *HTML 4.01*, *XHTML 1.0*, *XHTML Basic*, *XHTML 1.1*, *HTTP 1.1*, *MathML 2.0*, *CSS 2* a *SVG*. Implementácia žiadneho z týchto štandardov však nie je kompletná. Tento program nám nevyhovoval najmä preto, že niektoré matematické vzorce sa zobrazovali nekorektne, resp. niektoré symboly sa nezobrazovali vôbec a program obsahoval len niektoré základné matematické značky. Popis ďalších editorov možno nájsť na stránke konzorcia *W3C* [24].

3.1 Špecifikácia

Program (Obr. 3.1) na transformáciu textových súborov do *XML* súborov, spracovanie *XML* súborov, vkladanie matematických značiek jazyka *MathML*, vkladanie matematických symbolov a vzorcov s nasledujúcimi funkciami:



Obrázok 3.1:
Hlavné okno *XML Editor*.

- Súbor → Otvoriť
- Súbor → Otvoriť v prehliadači
- Súbor → Uložiť
- Súbor → Uložiť ako
- Súbor → Koniec
- Editovať → Transformovať
- Editovať → Prečíslovať

- Editovať → Exportovať do .csv
 - Postup
 - Pomoc
-
- Tlačidlo 1 – Základné značky
 - Tlačidlo 2 – Grécka abeceda
 - Tlačidlo 3 – Špeciálne znaky
 - Tlačidlo 4 – Šípky
 - Tlačidlo 5 – Matematické operátory

3.2 Popis funkcií

Funkcie Súbor → Otvoriť (CTRL+O), Súbor → Uložiť (CTRL+S), Súbor → Uložiť ako a Súbor → Koniec (ALT+F4) majú štandardnú funkcionality.

Funkcia Súbor → Otvoriť v prehliadači (CTRL+I) slúži na zobrazenie súboru, ktorý je otvorený v programe, v internetovom prehliadači.

Funkcia Editovať → Transformovať (CTRL+T) transformuje vybraný textový súbor do XML súboru. Vstupný textový súbor musí mať formát ako v príklade 3.2. Na základe tohto formátu program transformuje textový súbor do XML súboru, ktorý má štruktúru ako v príklade 3.3. Transformáciou prvého riadku textového súboru vznikne značka <oblast> s atribútom nazov (Príkl. 3.3, riadok 2), ktorý má príslušnú hodnotou z textového súboru (Príkl. 3.2, riadok 1). Transformáciou druhého riadku vznikne značka <definicia> s atribútom id (Príkl. 3.3, riadok 3). Transformáciou tretieho riadku vznikne značka <nazov> (Príkl. 3.3, riadok 4) s obsahom z textového súboru (Príkl. 3.2, riadok 3). Transformáciou štvrtého a piateho riadku vznikne značka <text> (Príkl. 3.3, riadok 5) s obsahom z textového súboru (Príkl. 3.2, riadok 4 – 5). Potom sa všetky párové značky ukončia a spracuje sa nasledujúca definícia. Koreňovou značkou XML súboru je <terminologia>.

```

1.  Matematická logika
2.  /
3.  (predmetové) premenné
4.  symboly x, y, z, w, x1, y1, z1, w1; množinu
5.  všetkých premenných budeme označovať znakom Var
6.  /

```

Príklad 3.2:

Formát vstupného textového súboru.

```

1.  <terminologia>
2.    <oblast nazov="Matematická logika">
3.      <definicia id="1">
4.        <nazov>(predmetové) premenné</nazov>
5.        <text> symboly x, y, z, w, x1, y1, z1, w1;
6.        množinu všetkých premenných budeme označovať
7.        znakom Var </text>
8.      </definicia>
9.    </oblast>
10. </terminologia>

```

Príklad 3.3:

Výstupný XML súbor.


Funkcia Editovať → Prečíslovať (CTRL+P) opraví všetky atribúty id elementov definícia. Postupne prechádza všetky id a vzostupne ich očísľuje, začína indexom 1.

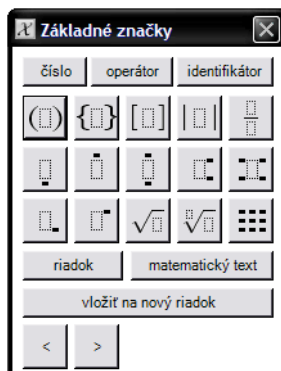
Funkcia Editovať → Exportovať do .csv exportuje súbor, ktorý je otvorený v programe, do súboru s príponou .csv. Výstupný súbor má nasledujúci formát, jeden riadok zodpovedá jednej definícii, hodnoty sú oddelené znakom \$ a zodpovedajú obsahu značiek <oblast>, <definicia id="">, <nazov> a <text>.

```
$oblast$číslo definície$názov definície$text definície$
```

V prípade terminológie z počítačovej grafiky obsahuje CSV súbor aj položku \$anglický termín\$, ktorá zodpovedá obsahu značky <english>.


Funkcie Postup a Pomoc otvoria internetové stránky s príslušným obsahom. V prípade prvej funkcie je to sekcia Postup [26] a v druhom prípade je to sekcia XML Editor [27].

Tlačidlo  otvorí formulár Základné značky (Obr. 3.4). Nachádzajú sa tu tlačidlá s matematickými značkami na vytváranie vzorcov.



Obrázok 3.4:


Tlačidlá na vkladanie matematických značiek.

Tlačidlo  otvorí formulár Grécka abeceda (Obr. 3.5). Tento formulár slúži na vkladanie veľkých a malých písmen gréckej abecedy.



Obrázok 3.5:


Tlačidlá na vkladanie gréckej abecedy.

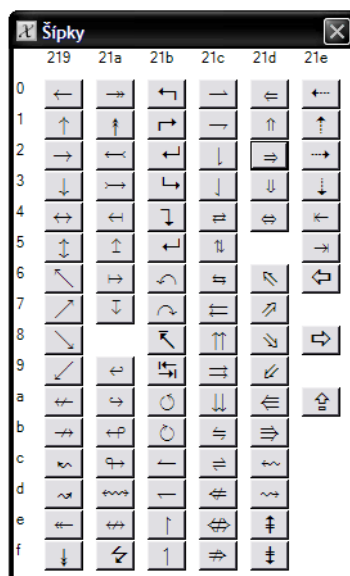
Tlačidlo  otvorí formulár Špeciálne znaky (Obr. 3.6), ktorý slúži na vkladanie špeciálnych znakov.



Obrázok 3.6:


Tlačidlá na vkladanie špeciálnych znakov.

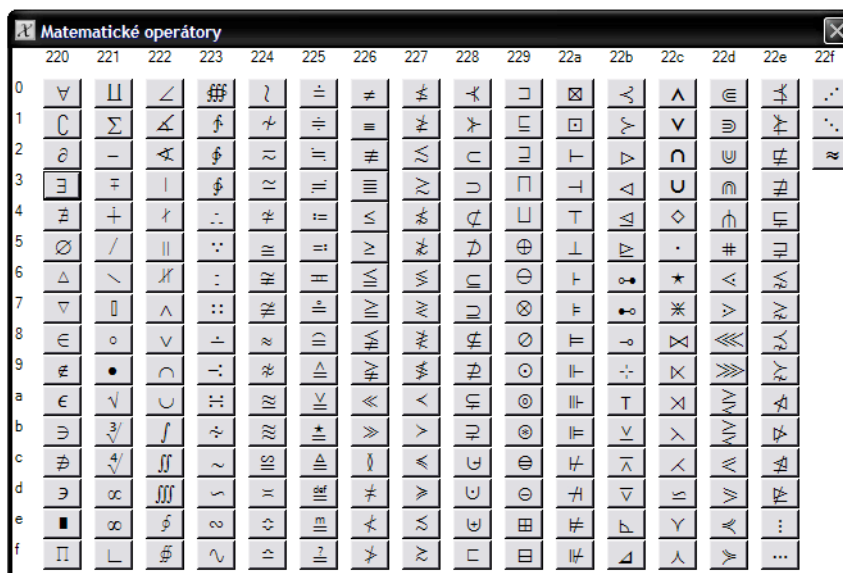
Tlačidlo  otvorí formulár Šípky (Obr. 3.7). Tento formulár slúži na vkladanie rôznych šípok.



Obrázok 3.7:

Tlačidlá na vkladanie šípok.

Tlačidlo  otvorí formulár Matematické operátory (Obr. 3.8), ktorý slúži na vkladanie matematických operátorov.



Obrázok 3.8:

Tlačidlá na vkladanie matematických operátorov.

Všetky formuláre obsahujú názvy stĺpcov a riadkov, zložením ktorých sa dá určiť kód znaku *unicode*. Napríklad, znak Σ má vo formulári Matematické operátory kód 2211. Takto získaný kód sa zapíše do *XML* súboru takto:

```
<mo>&#x2211;</mo>
```

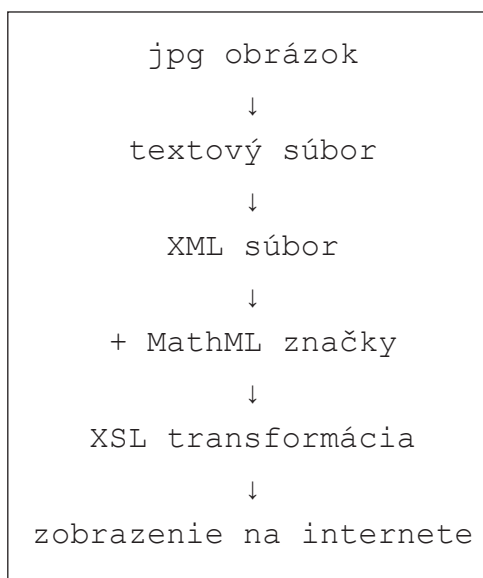
4 Postup práce

O svojom návrhu môžeš uvažovať akokoľvek, ostatní ho posúdia ináč.

CHISHOLMOV TRETÍ ZÁKON

4.1 Postup

Postup našej práce sme rozdelili do niekoľkých na seba nadväzujúcich krokov (Obr. 4.1). Vstupom boli *jpg* obrázky naskenovaných textov, texty [14] boli napísané na písacom stroji. Tieto obrázky sme ďalej pomocou OCR softvéru *Abby FineReader* transformovali do textových súborov. Časť obrázkov bola vo výbornej kvalite (Obr. B.1, príloha B) a preto nebol problém ani z ich rozpoznaním. Problémy sme mali najmä s obrázkami v zlej kvalite a s tými, ktoré obsahovali ručne vpisované poznámky (Obr. B.2, príloha B). Tieto obrázky sme preto najprv upravili a až následne potom sme ich rozpoznávali. Pôvodne táto časť mala byť vynechaná, pretože materiály mali byť dodané už rozpoznané. Omeškal sa však vývoj softvéru.



Obrázok 4.1:
Postup práce.

V ďalšom kroku sme rozpoznané texty upravili do jednoduchého formátu (Príkl. 4.2).

```
1.  Matematická logika
2.  /
3.  (predmetové) premenné
4.  symboly  $x, y, z, w, x_1, y_1, z_1, w_1$ ; množinu
5.  všetkých premenných budeme označovať znakom Var
6.  + preformulovať
7.  /
8.  logické symboly [=operátory, spojky]
9.  symboly typu: (negácia), (disjunkcia), (konjunkcia),
10. (implikácia), (ekvivalencia), (všeobecný kvantifikátor),
11. (existenčný kvantifikátor)
12. /
13. ...
```

Príklad 4.2:

Formát textového súboru.

Prvý riadok obsahuje názov oblasti, z ktorej je terminológia (Príkl. 4.2, riadok 1). Ďalej nasleduje znak /, ktorý označuje začiatok definície (Príkl. 4.2, riadok 2). Názov definície sa nachádza na ďalšom riadku. Potom nasleduje text definície, ktorý môže byť na ľubovoľnom počte riadkov (Príkl. 4.2, riadok 4 – 5). Znak + označuje poznámku, ktorá nie je povinná (Príkl. 4.2, riadok 6). Táto poznámka slúži na poznamenanie si nejakého postrehu k definícii, nachádza sa v zdrojovom kóde *XML* súboru, ale v internetovom prehliadači sa nezobrazí. Potom opäť nasleduje znak / označujúci novú definíciu (Príkl. 4.2, riadok 7). Súboru sme uložili s kódovaním *UTF-8*.

Takto upravené textové súboru boli vstupom pre transformáciu do *XML* súborov. Na transformáciu sme použili *XML Editor* a jeho funkciu Editovať → Transformovať. Po transformácii vznikli súboru so štruktúrou ako v príklade 4.3. Viac o transformácii pomocou *XML Editoru* sa nachádza v kapitole 3.2 na strane 25.

```

<?xml version="1.0" encoding="utf-8"?>
<?xml-stylesheet type="text/xsl" href="trans/trans.xsl"?>
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.1 plus MathML
2.0//EN" "http://www.w3.org/TR/MathML2/dtd/xhtml-math11f.dtd"
[<!ENTITY mathml "http://www.w3.org/1998/Math/MathML">]>
<terminologia>
  <oblast nazov="Matematická logika">
    <definicia id="1">
      <nazov>
        <math xmlns="&mathml;">
          (predmetové) premenné
        </math>
      </nazov>
      <text>
        <math xmlns="&mathml;">
          symboly  $x, y, z, w, x_1, y_1, z_1, w_1$ ;
          množinu všetkých premenných budeme
          označovať znakom  $\text{Var}$ 
        </math>
      </text>
    </definicia>
  </oblast>
</terminologia>

```

Príklad 4.3:

XML súbor transformovaný z textového súboru.

Posledným krokom bola úprava *XML* súborov. V týchto súboroch bolo potrebné upraviť najmä matematické symboly a vzorce, ktoré boli nekorektne rozpoznané pri transformácii do textu. Na úpravu sme použili *XML Editor*. Výsledný zdrojový kód je zobrazený v príklade 4.4.

```

<?xml version="1.0" encoding="utf-8"?>
<?xml-stylesheet type="text/xsl" href="trans/trans.xsl"?>
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.1 plus MathML
2.0//EN" "http://www.w3.org/TR/MathML2/dtd/xhtml-math11f.dtd"
[<!ENTITY mathml "http://www.w3.org/1998/Math/MathML">]>
<terminologia>
  <oblast nazov="Matematická logika">
    <definicia id="1">
      <nazov>(predmetové) premenné</nazov>
      <text>
        <math xmlns="&mathml;">
          <mtext>symboly</mtext>

```



```

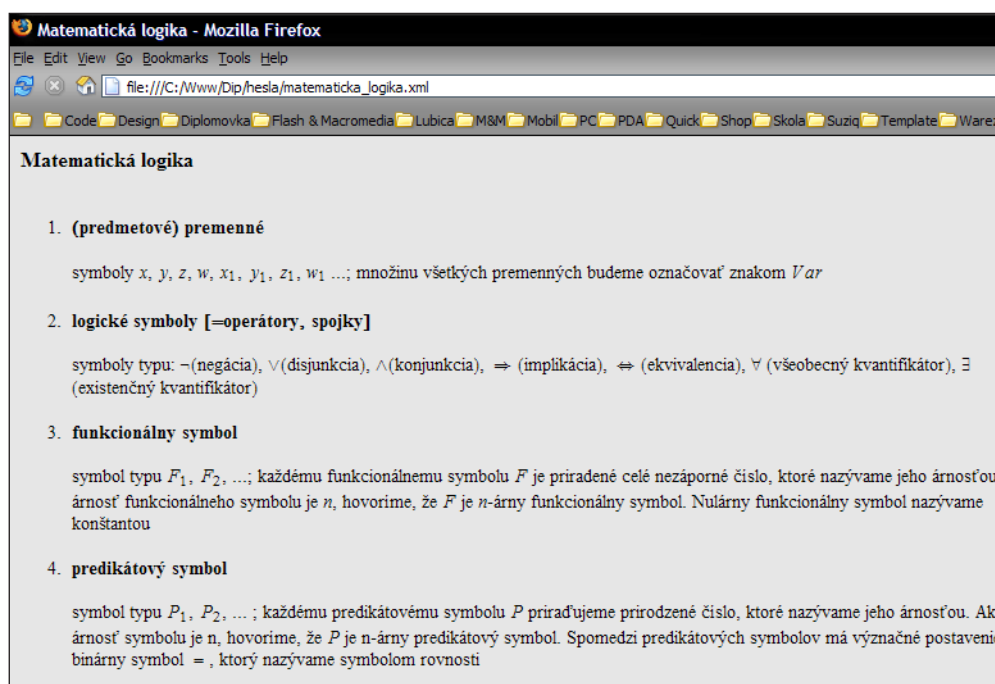
<mi>x</mi>
<mtext>,</mtext>
<mi>y</mi>
<mtext>,</mtext>
<mi>z</mi>
<mtext>,</mtext>
<mi>w</mi>
<mtext>,</mtext>
<msub>
  <mi>x</mi>
  <mn>1</mn>
</msub>
<mo>,</mo>
<msub>
  <mi>y</mi>
  <mn>1</mn>
</msub>
<mo>,</mo>
<msub>
  <mi>z</mi>
  <mn>1</mn>
</msub>
<mo>,</mo>
<msub>
  <mi>w</mi>
  <mn>1</mn>
</msub>
<mi>...</mi>
<mtext>množinu všetkých premenných
  budeme označovať znakom
</mtext>
<mrow>
  <mi>V</mi>
  <mi>a</mi>
  <mi>r</mi>
</mrow>
</math>
</text>
</definicia>
</oblast>
</terminologia>

```

Príklad 4.4:
XML súbor po úprave.

Takto upravené *XML* súbory boli konečným výstupom a mohli sme si ich zobrazíť v internetovom prehliadači *Mozilla Firefox*, ktorý podporuje zobrazovanie matematických symbolov a vzorcov (Obr. 4.5).

Na internetovej stránke [26] k tejto práci v sekcii *Postup* môžete nájsť tento postup spolu s použitím *XML Editor*a a s príkladmi.



Obrázok 4.5:

Zobrazenie *XML* súboru na internete.

4.2 *XML* súbory

- algebraicka_geometria.xml
- algebraicke_struktury.xml
- aritmetika_a_teorija_cisel.xml
- deskriptivna_geometria.xml
- geometria_vseobecne_pojmy.xml
- kombinatorika.xml
- linearna_algebra.xml
- matematicka_logika.xml
- numericka_matematika.xml

- pocitacova_grafika.xml
- pravdepodobnost_a_statistika.xml
- projektivna_geometria.xml
- teoria_grafov.xml
- teoria_kategorii.xml
- teoria_mnozín.xml
- topologia.xml
- usporiadane_algebraicke_struktury.xml

Výstupných *XML* súborov je 17 (Príloha C), pre každú oblasť sme vytvorili jeden súbor. Termíny z matematickej oblasti sme spracovali z poskytnutých materiálov [14], [15] a [16] a všetky príslušné súbory majú rovnakú štruktúru (Príkl. 4.6). Na ich zobrazenie na internete sa použili rovnaký transformačný súbor `trans.xsl`.

```
<terminologia>
  <oblast nazov=" ">
    <definicia id=" ">
      <nazov>
        <math xmlns="&mathml;"> </math>
      </nazov>
      <text>
        <math xmlns="&mathml;"></math>
      </text>
    </definicia>
  </oblast>
</terminologia>
```

Príklad 4.6:
Štruktúra *XML* súborov.

Koreňovou značkou *XML* súboru je `<terminologia>`, potom nasleduje značka `<oblast nazov=" ">`, ktorej atribút `nazov` určuje oblasť z ktorej termíny pochádzajú. Značka `<definicia id=" ">` obsahuje vždy jednu definíciu, ktorej obsah je rozdelený medzi dve značky `<nazov>`, ktorá obsahuje názov definície a `<text>`, ktorá obsahuje text definície. Atribút `id` určuje číslo definície. V každej značke `<nazov>` a `<text>` sa vyskytuje značka `<math xmlns="&mathml;">`, ktorá obsahuje entitu `&mathml;`. Táto entita odkazuje na menný priestor (*namespace*) *MathML* značiek [29] a

nachádza sa tam preto, aby boli zobrazené všetky matematické symboly a vzorce korektne. Súbor na transformáciu XML súboru sa priradí pomocou inštrukcie na spracovanie takto:

```
<?xml-stylesheet type="text/xsl" href="trans/trans.xsl"?>
```

Transformačný súbor obsahuje postupnosť XSL príkazov na spracovanie jednotlivých značiek z XML súboru. Tieto značky formátuje do podoby čitateľnej na internete. Určuje druh, veľkosť, štýl, farbu písma, rozmiestnenie elementov a tiež určuje celkové vizuálne zobrazenie stránky.

Súbor obsahujúci termíny z počítačovej grafiky má podobnú štruktúru (Príkl. 4.7) a na transformáciu sa používajú súbory pg.xsl a pg_en.xsl (Príloha C).

```
<terminologia>
  <zdroj1> </zdroj1>
  <zdroj2> </zdroj2>
  <zdroj3> </zdroj3>
  <zdroj4> </zdroj4>
  <zdroj5> </zdroj5>
  <zdroj6> </zdroj6>
  <zdroj7> </zdroj7>
  <zdroj8> </zdroj8>

  <triedenie-sk>sk</triedenie-sk>
  <triedenie-en>en</triedenie-en>

  <oblast nazov="Počítačová grafika">
    <definicia id=" ">
      <nazov> </nazov>
      <text1> </text1>
      <text2> </text2>
      ...
      <text6> </text6>
      <english> </english>
    </definicia>
  </oblast>
</terminologia>
```

Príklad 4.7:

Štruktúra súboru pocitacova_grafika.xml.

Koreňovou značkou je <terminologia>, potom nasledujú značky <zdrojX> s číselným označením. Každá značka obsahuje jeden zo zdrojov [1], [2], [3], [4], [5], [6], [7], [8] z ktorej bola terminológia spracovaná. Značky <triedenie-sk> a <triedenie-en> slúžia na určenie usporiadanie termínov podľa slovenského názvu alebo podľa anglického názvu. Značka <oblast nazov=" "> obsahuje atribút nazov, ktorý určuje oblasť z ktorej termíny pochádzajú. Potom nasleduje značka <definicia id=" ">, pre každú definíciu jedna. Obsah definície je rozdelený medzi ďalšie značky. Značka <nazov> obsahuje názov definície, <textX> obsahuje text definície, podľa čísla v značke sa dá určiť zdroj definície. Značka <english> obsahuje anglický názov definície.

V XML súbore sa ešte vyskytujú značky <odkaz href=" "> a . Prvá značka slúži na vkladanie odkazov do textu, atribút href musí obsahovať id číslo niektorej z definícií. Druhá značka slúži na vkladanie obrázkov do textu. Atribút src musí obsahovať zdroj obrázku, atribút popis slúži na krátky popis obrázku, ktorý sa zobrazí pod obrázkom v internetovom prehliadači. Atribúty sirka a dlzka definujú rozmery obrázku v pixloch a odkaz obsahuje adresu odkiaľ bol obrázok získaný.

Na transformáciu sme vytvorili dva XSL súbory pg.xsl a pg_en.xsl. Obidva zobrazujú XML súbor rovnako, prvý však utriedi termíny podľa slovenského názvu, druhý podľa anglického názvu. Na triedenie sme použili XSL príkaz <xsl:sort select=" " />, ktorého atribút select určuje značku, podľa ktorej sa aplikuje triedenie. V prípade prvého transformačného súboru sme použili značku <nazov>, v druhom prípade značku <english>. Na číslovanie definícií sme použili XSL príkaz <xsl:number value=" " format=" " />, ktorého atribút value určuje hodnoty a format určuje formát číslovania. Na určovanie číselných hodnôt atribútu value sme použili funkciu position(), ktorá určí číslo na základe výskytu značky <definicia> v XML súbore. Hodnotu pre formát číslovania sme použili format="1. ". Na uľahčenie práce s XML súborom sa po transformácii v texte nachádza aj abecedou indexovaná navigácia.

Transformačné súbory sme sa snažili vytvoriť tak, aby bolo možné *XML* súbory neustále dopĺňať o nové definície. Aby nebolo potrebné ich manuálne triediť a prečíslovať. Na správnu štruktúru a validitu *XML* súborov sme vytvorili aj *DTD (Document Type Definition)*. Ďalšou výhodou je, že sme oddelili obsah *XML* súborov od ich formátu použitím transformačného súboru. T. j. *XML* súbory sa dajú opakovane použiť s ľubovoľným transformačným súborom. To znamená, že v budúcnosti sa na ne môže aplikovať ľubovoľná iná transformácia, napr. na vytvorenie knižnej podoby.

5 Internetová stránka

Najjednoduchšie myšlienky sa najzložitejšie vyjadrujú.

MALEKOV ZÁKON

Najjednoduchší spôsob ako zobrazíť matematický vzorec na internete je vytvoriť ho v nejakom editore a uložiť ako obrázok. Nevýhodou tohto riešenia je, že pri nejakej zmene je nutné ho opäť vytvoriť, editovať a uložiť. Naším primárnym cieľom bolo ale zobrazenie matematickej terminológie a terminológie z počítačovej grafiky na internete pomocou jazyka *XML* a *MathML*. Podpora týchto jazykov, resp. štandardov je najlepšia v internetovom prehliadači *Mozilla Firefox*.

5.1 Štruktúra

Stránka sme vytvorili pomocou štandardov *XHTML 1.0 Transitional*, *CSS*, skriptovacieho jazyka *PHP* a použili sme databázu *MySQL*. Na stránke sa nachádzajú 4 základné sekcie: *Úvod*, *O projekte*, *Pomoc*, *Kontakt*. Obsahovú časť stránky sme rozdelili do nasledujúcich 12 sekcií:

- *Termíny* – prehľadová tabuľka terminológie
- *Pridať termín* – formulár na pridanie termínu
- *Hľadať termín* – formulár na vyhľadávanie termínov
- *Postup* – návod ako postupovať pri spracovaní terminológie
- *XML Editor* – informácie o *XML Editore*
- *Materiály* – zoznam materiálov vytvorených počas práce
- *EN – SK slovník* – Slovník pojmov CSN, anglicko – slovenský podľa ERTL, J. – FERKO, A. 1992
- *Malý spisovný slovník* – z redakčnej praxe *PC REVUE*
- *Anketa* – anketa na analýzu odozvy matematickej komunity
- *Kniha návštev* – odkazy
- *Informácie* – informácie o stránke

- *Odkazy* – odkazy na programy a potrebné informácie
- *História stránky* – chronologická história zmien na stránke

5.2 Obsah

Sekcia *Termíny* obsahuje prehľadovú tabuľku (Tabuľka 5.1), ktorá sumarizuje počet všetkých termínov, ktoré sme spracovali.

Názov	Počet hesiel	Heslár
• Terminológia prístupná na heslo:		
Algebraická geometria	176	odkaz
Algebraické štruktúry	326	odkaz
Aritmetika a teória čísel	231	odkaz
Deskriptívna geometria	161	odkaz
Geometria - všeobecné pojmy	289	odkaz
Kombinatorika	88	odkaz
Lineárna algebra	53	odkaz
Matematická logika	66	odkaz
Numerická matematika	269	odkaz
Pravdepodobnosť a štatistika	235	odkaz
Projektívna geometria	202	odkaz
Teória grafov	179	odkaz
Teória kategórií	80	odkaz
Teória množín	155	odkaz
Topológia	234	odkaz
Usporiadané algebraické štruktúry a topologická algebra	102	odkaz
• Terminológia voľne prístupná:		
Počítačová grafika	393	odkaz
Spolu	3239	

Tabuľka 5.1:

Zoznam XML súborov s počtom termínov.

Prístup k matematickej terminológii je na heslo, pretože je autorským dielom kolektívu pod vedením prof. Jána Čizmára, autori nesúhlasili zatiaľ so zverejnením a časti tejto práce nie sú dokončené. Všetky XML súbory sa

nachádzajú v zaheslovanom adresári. Terminológia z počítačovej grafiky je voľne prístupná pre všetkých používateľov internetu.

Obrázok 5.2:

Formulár na vkladanie termínov do databázy.

Sekcia *Pridať termín* obsahuje formulár na vloženie termínov do databázy, ktoré používatelia nenašli v spracovanej terminológii (Obr. 5.2). Vložené dáta sa ukladajú do tabuliek `matematika`, `grafika`, `programovanie`, `internet` a `ina_oblast` podľa toho, ktorú oblasť používateľ vybral. Každá z týchto tabuliek má štruktúru ako tabuľka 5.3.

Názov poľa	Typ poľa	Akcia	Popis
id	int	auto_increment	primárny identifikátor
termin	text		pole termín z formulára
popis	text		pole popis z formulára
oblast	text		pole iná oblasť z formulára
zdroj	text		pole autor/zdroj z formulára

Tabuľka 5.3:

Štruktúra tabuliek na ukladanie termínov z formulára (Obr. 5.2).

Sekcia *Hľadať termín* obsahuje formulár na hľadanie termínov zo spracovanej matematickej terminológie a terminológie z počítačovej grafiky (Obr. 5.4).

Termín:

Presný výraz

Oblasť:

- Algebraická geometria
- Algebraické štruktúry
- Aritmetika a teória čísiel
- Deskriptívna geometria
- Geometria - všeobecné pojmy
- Kombinatorika
- Lineárna algebra
- Matematická analýza
- Matematická logika
- Numerická matematika
- Pravdepodobnosť a štatistika
- Projektívna geometria
- Teória grafov
- Teória kategórií
- Teória množín
- Topológia
- Usporiadané algebraické štruktúry a topologická algebra
- Počítačová grafika

Obrázok 5.4:

Formulár na vyhľadávanie termínov.

Na vyhľadávanie sme použili databázu *MySQL*. Do tabuľky *termíny* sme uložili všetky *XML* súbory spracovanej matematickej terminológie a terminológie z počítačovej grafiky. Súbory sme však najprv pomocou *XML Editora* transformovali do *CSV* súborov. Popis tejto transformácie nájdete v kapitole 3.2 na strane 26. Potom sme ich prostredníctvom internetovej aplikácie *phpMyAdmin 2.6.4* a príkazom *Vložiť* textové súbory do tabuľky vložili do tabuľky *termíny*. Táto tabuľka obsahuje termíny zo všetkých matematických oblastí a aj termíny z počítačovej grafiky a má štruktúru ako tabuľka 5.5.

Názov poľa	Typ poľa	Akcia	Popis
id	int	auto_increment	primárny identifikátor
oblast	text		oblasť z ktorej je definícia
id_def	text		číslo definície

nazov	text		názov definície
text	text		text definície
eng	text		anglický termín

Tabuľka 5.5:
Štruktúra tabuľky terminy.

Používateľ môže vo formulári zaškrtnúť položku *Presný výraz*. Ak zadá používateľ napríklad slovo *grupa*, tak v tomto prípade budú nájdené iba definície s presným termínom *grupa*. V opačnom prípade budú všetky nájdené definície obsahovať aj termíny ako napríklad *podrupa*, *grupoid*, *pologrupa*, teda aj termíny, kde *grupa* je podslovo. Každá nájdená definícia obsahuje odkaz na súbor, v ktorom sa nachádza. Používateľ si tak môže pozrieť všetky definície z príslušnej oblasti.

Sekcia *Postup* obsahuje návod ako postupovať pri spracovávaní terminológie. Tento postup sme vytvorili pre katedru, ktorej úlohou je spracovať terminológiu z matematickej analýzy. Návod má 4 kroky:

1. *Inštalácia*
2. *Transformácia jpg → textový súbor, textový súbor → XML súbor*
3. *Vkladanie matematických symbolov a vzorcov*
4. *Zobrazenie výstupného súboru*

Prvý krok obsahuje stručný návod na inštaláciu potrebných programov a matematického písma. Druhý krok popisuje jednotlivé kroky transformácie vstupných materiálov do výstupných XML súborov. Tretí krok uvádza ako použiť *XML Editor* na vkladanie matematických symbolov a vzorcov. Štvrtý krok popisuje postup zobrazenia XML súborov v internetovom prehliadači s použitím pripojenia transformačných XSL súborov.

Sekcia *XML Editor* obsahuje históriu vývoja tohto programu, popis zmien v jednotlivých verziách a odkazy, kde si môžete tieto verzie stiahnuť.

Sekcia *Materiály* obsahuje prezentácie, ktoré vznikli počas tejto práce a pdf texty k tejto práci, ktoré si môžete stiahnuť. Ďalej sa tu nachádza zoznam spracovaných materiálov.

Sekcia *EN – SK slovník* obsahuje anglicko – slovenské termíny z počítačovej grafiky [7].

Sekcia *Malý slovník* obsahuje tabuľku s dvomi stĺpcami. V prvom stĺpci s názvom *Nesprávne* sa nachádzajú nespisovné výrazy. V druhom stĺpci s názvom *Správne* sa nachádzajú spisovné výrazy. Slovník pochádza z redakčnej praxe *PC REVUE*.

Sekcia *Anketa* obsahuje formulár s ôsmimi otázkami, ktorý sa ukladá do databázy. Odpovede respondentov sa nachádzajú v tabuľke anketa, ktorá má štruktúru ako tabuľka 5.6.

Názov poľa	Typ poľa	Akcia	Popis
id	int	auto_increment	primárny identifikátor
otazka1	text		odpoveď na otázku č.1
...	
otazka8	text		odpoveď na otázku č.2

Tabuľka 5.6:
Štruktúra tabuľky anketa.

Anketu sme vytvorili na získanie informácií o používaní internetovej stránky, ktorú sme vytvorili v rámci našej práce. Viac o analýze výsledkov sa dozviete v kapitole 5.3 na strane 45.

Sekcia *Kniha návštev* obsahuje formulár na vkladanie odkazov na stránku. Odkazy sú ukladajú do tabuľky kniha, ktorá ma štruktúru ako tabuľka 5.7.

Názov poľa	Typ poľa	Akcia	Popis
id	int	auto_increment	primárny identifikátor
datum	text		dátum vloženia odkazu
meno	text		pole meno z formulára
email	text		pole email z formulára
web	text		pole web z formulára
odkaz	text		pole odkaz z formulára

Tabuľka 5.7:
Štruktúra tabuľky kniha.

Sekcia *Informácie* obsahuje informácie ako so stránkou pracovať, ako si ju prispôbiť a ako získať heslo na prístup k terminológii.

V sekcii *História stránky* je uvedená chronologická história zmien na stránke od jej vzniku.

5.3 Analýza výsledkov ankety

Jedným z cieľov našej práce bolo otestovanie výsledkov práce v matematickej komunite. Ku dňu 10. 05. 2006 do ankety prispelo 37 respondentov. Anketa nám umožnila identifikovať a odstrániť niektoré drobné nezrovnalosti. Pýtali sme sa na 8 otázok. Všetky nasledujúce odpovede sú originálne, preto môžu obsahovať expresívne, resp. nespisovné výrazy.

Otázka č. 1: Aký je Váš celkový dojem zo stránky? Odpovede na prvú otázku obsahovali prevažne nasledujúce slová:

pozitívne:

- jednoduchá
- príjemná
- farebne zladená
- ucelená
- slušná
- prehľadná
- nevtieravá

negatívne:

- príliš konzervatívna
- málo farebná
- fádna

Na základe týchto odpovedí sme sa rozhodli nerobiť žiadne zmeny týkajúce sa dizajnu stránky.

Otázka č. 2: Aký je Váš názor na obsah? Odpovede na druhú otázku obsahovali nasledujúce postrehy:

pozitívne:

- zaujímavá téma
- zaujímavý projekt
- našla som veci, ktoré sa zídu

negatívne:

- málo slov
- neprehľadný
- mohol by mať väčší záber

- celkom slušný slovník
- ešte som nevidel nič podobného, zhrnutie na jednej stránke, krása
- zodpovedá zámeru
- potrebná vec
- počítačová grafika sa zíde
- vhodné pre toho čo rýchlo niečo hľadá
- spokojnosť
- super pre tých čo potrebujú si rýchlo nájsť nejakú definíciu
- pozrel som len grafiku a tá je celkom dobrá
- obsah je v poriadku
- mohlo by byť viac

Väčšina negatívnych odpovedí obsahovala názor, že obsahu by mohlo byť viac, resp. že počet termínov je nedostatočný. Našou úlohou bolo spracovať poskytnuté materiály v rozsahu asi 660 strán formátu A4. Všetky termíny nachádzajúce sa v týchto materiáloch boli kompletne spracované. Samozrejme, že pridávanie nových termínov resp. ich aktualizácia v budúcnosti je možná a bude sa realizovať.

Otázka č. 3: Zobrazujú sa Vám všetky informácie korektne (matematické symboly, diakritika, obrázky)? Odpovede na túto otázku boli zhodné, resp. všetkým respondentom sa stránka zobrazovala korektne. Nevyhnutnosťou na správne zobrazenie najmä matematických symbolov a vzorcov je použitie internetového prehliadača *Mozilla Firefox* a inštalácia matematického písma. Návod na inštaláciu a aj potrebné inštalačné balíky sa nachádzajú na internetovej stránke v sekcii *Postup* [26]. Jedna odpoveď obsahovala názor, že internetová stránka [25] funguje aj v internetovom prehliadači *Internet Explorer 7 Beta*, čo je pozitívna informácia do budúcnosti, pretože podpora jazyka *MathML* v tomto prehliadači je momentálne nedostačujúca.

Otázka č. 4: Máte pripomienky k funkčnosti? Odpovede na štvrtú otázku boli nasledovné:

pozitívne:

- neviem žiadne
- mne sa to zdá byť celkom funkčné
- nie

negatívne:

- nejde v IE
- pridal by som referenciu z jedného termínu na iný, ak sa ten druhý používa v definícii toho prvého
- keď som si zapla slovník počítačovej grafiky, tak nefungovali horné záložky ako „Domov”... všetky
- pri nájdení výsledkov sa zbytočne po kliknutí na linku otvára nové okno

Na základe týchto odpovedí sme zrušili na internetovej stránke [25] otváranie odkazov v novom okne, pretože to je zbytočné a len to spomaľuje prácu. Opravili sme všetky nefungujúce hypertextové odkazy. Referencie, resp. odkazy na termíny, ktoré sa nachádzajú v definícii existujú a sú funkčné v súbore `pocitacova_grafika.xml`.

Otázka č. 5: Uvítame Vaše návrhy na zlepšenie. Odpovede na piatu otázku boli nasledovné:

pozitívne:

- zatiaľ žiadne
- nemám
- nič mi nenapadá
- nemám žiadne

negatívne:

- viac termínov
- zvalidnite si stránku a css
- väčšie textarea bloky v ankete
- tá farebnosť, ale nie prehnaná, no podľa mňa sa to dosť blbo číta
- terminológia (pre PG) nie je konzistentná (viď. hidden line, hidden line removal)
- rozšíriť množstvo definícii
- pri vyhľadávaní chýba možnosť „označ všetky kategórie”
- pridávať stále heslá, aby bol obsah stále väčší
- textové vyhľadávanie

Na základe týchto odpovedí sme na internetovej stránke [25] urobili viacero zmien. Stránku sme kompletne validovali pomocou služby *Markup Validation Service*, ktorá sa nachádza na stránke konzorcia W3C [28]. V sekcii *Hľadať termín* sme pridali tlačidlá *Označiť všetko* a *Odznačiť všetko* a tiež sme pridali formulár na textové vyhľadávanie v rámci stránky, ktorý sa nachádza v sekciami *Úvod*, *O projekte*, *Pomoc*, *Kontakt* a *Informácie*. V prípade nezrovnalostí, týkajúcich sa nekonzistentnosti termínov bude vykonaná odborná revízia.

Otázka č. 6: Poznáte podobný projekt (aj v inej oblasti) na internete? Ak áno, poskytnite prosím jeho URL. Pomôže nám porovnať a/alebo vylepšiť tento projekt. Väčšina respondentov odpovedala, že nepozná žiadny podobný projekt. Z príbuzných projektov bola najčastejšie spomenutá Wikipedia [22] a Slovník slovenského jazyka [30].

Otázka č. 7: Aká je Vaša kvalifikácia (študent grafiky, študent, gymnazista, učiteľ, iné)? Do ankety prispeli najmä študenti (48,64%), študenti grafiky (45,94%), učitelia a IT odborníci (5,40%).

Otázka č. 8: Akú otázku by ste pridali do ankety? Väčšina respondentov odpovedala, že žiadnu otázku by nepridala. Návrhy na ďalšiu otázku boli nasledovné:

- aký systém daný užívateľ užíva: Linux, Win, MacOS
- podrobné spracovanie akej témy by vás potešilo/pomohlo vám?
- čokoľvek čo mi chcete napísať k danej stránke...
- využijete túto stránku?
- túto by som odstránil

Vzhľadom ku krátkemu času sprístupnenia možno výsledky ankety hodnotiť ako nad očakávanie početne a dosť reprezentatívne, lebo odozva prišla z vysoko kvalifikovanej komunity. Celkové zhodnotenie je zobrazené v tabuľke 5.8.

Číslo otázky	Počet odpovedí	Odpovede na otázku	Hodnota v %
1.	37	28 – prevažne kladných	75,67
2.	37	26 – prevažne kladných	70,27
3.	37	31 – prevažne áno	83,78
4.	37	20 – prevažne nie	54,05
5.	37	26 – prevažne nie	70,27
6.	37	27 – nepoznám	72,97
		8 – Wikipedia	21,62
		2 – iné	5,40
7.	37	17 – študent grafiky	45,94
		18 – študent	48,64
		2 – iné	5,40

Tabuľka 5.8:
Štatistický prehľad odpovedí z ankety.

Záver

Nech sa vyjadrite akokoľvek presne, každý to pochopí inakšie.

ZÁKON VÝZNAMOVÉHO ROZPTYLU

Cieľom našej práce bolo spracovať rukopis matematickej terminológie, navrhnúť formu publikácie, implementovať a otestovať v matematickej komunite. Hlavným prínosom našej práce je spracovaných 660 strán matematickej terminológie do 16 samostatných súborov, kde jeden súbor obsahuje termíny z jednej matematickej oblasti. Vytvorili sme aj jeden súbor obsahujúci termíny z počítačovej grafiky. Spolu 17 súborov vo formáte *XML*. Pri spracovávaní sme vytvorili originálny postup, ktorý sa dá použiť na spracovanie ďalších nových materiálov, resp. na aktualizáciu už vytvorených. Tento postup zahŕňa aj softvérový program *XML Editor*, ktorý sme vytvorili na editáciu *XML* súborov a na jednoduchšie vytváranie matematických vzorcov a vkladanie matematických symbolov.

Všetky súbory, obsahujúce dokopy 3239 termínov, sme publikovali online na internetovej stránke, kde si môže používateľ termíny nielen prezerať, ale aj pomocou formulára v nich vyhľadávať, t. j. obohatili sme funkcionality o interaktivitu. Navyše termíny z počítačovej grafiky obsahujú ilustračné obrázky a hypertextové odkazy na termíny vyskytujúce sa v definícii jednotlivých termínov.

Ďalší prínos našej práce sme testovali v rámci ankety, v ktorej sme sa respondentov pýtali na ich názor na obsah, celkové spracovanie a zobrazenie obsahu. Výsledky ankety indikujú prevažne pozitívne prijatie výsledkov projektu cieľovými skupinami. Na základe odpovedí sme mohli našu prácu ešte vylepšiť o ďalšiu funkcionality. Pridali sme fulltextové vyhľadávanie. Stránku sme kompletne validovali, t. j. spĺňa štandardy *XHTML 1.0 Transitional* a *XHTML 1.1 plus MathML 2.0*. V rámci uľahčenia práce s obsahom sme pridali abecedou indexovanú navigáciu. Celkovo anketa preukázala pozitívny ohlas v matematickej komunite.

Samozrejme terminológia nie je finálna a bude podliehať odbornej revízii v rámci príslušných katedier. Použitie *XML* štandardu dáva do budúcnosti neobmedzené možnosti využitia spracovaných materiálov.

Po získaní súhlasu autorov, resp. ich dedičov sa výsledný rukopis bude publikovať v elektronickej alebo knižnej podobe.

Literatúra

- [1] ISO/IEC 2382-13 – Rukopis slovenského prekladu (A. Ferko)
- [2] J. Ragan, *Anglicko – slovenský slovník výpočtovej techniky*, SPN, 1998
- [3] Výskumné výpočtové stredisko, *Terminologické listy – Základné termíny z oblasti počítačovej grafiky*, 1979
- [4] E. Ružický, A. Ferko, *Počítačová grafika*, Sapienta, 1995
- [5] J. Žára, B. Beneš, J. Sochor, P. Felkel, *Moderní počítačová grafika*, Computer Press, 2004
- [6] V. Medek, J. Zámožník, *Osobný počítač a geometria*, Alfa, 1991
- [7] J. Ertl, A. Ferko, *Normalizované grafické systémy*, Univerzita Komenského Bratislava, 1993
- [8] J. Vaský, M. Klačo, E. Nemlaha, *Grafické spracovanie údajov*, Slovenská technická univerzita v Bratislave, 2000
- [9] M. Urbanová, M. Oralková, *Angličtina vo výpočtovej technike*, Alfa, 1989
- [10] G. Žetkuliak, *Predmet a štruktúra informatiky – Terminológia informatiky*, Bratislava 1971
- [11] Z. Jurčocková, *Terminológia*, Centrum VTI SR Bratislava, 2002
- [12] J. Čižmár, *Tradičné a aktuálne problémy slovenskej matematickej terminológie, špeciálne geometrickej*, SCG' 2004 str. 32 – 39

- [13] J. Čižmár, *Korektná terminológia – súčasť odbornej kvalifikácie na tvorbu učebníc*, Úloha učebnice vo vyučovaní matematiky, Nitra 2004 str. 14 – 18
- [14] Igor Bock, Juraj Bosák, Miloš Božek, Juraj Činčura, Jan Čižmár, Arnold Dávid, Jozef Dravecký, Eva Gedeonová, Michal Grajcar, Milan Hamala, Teodor Hlaváč, Mirko Horňák, Ivan Korec, Anton Legéň, Václav Medek, Milan Medveď, Tibor Neubraunn, Ján Plesník, Štefan Porubský, Pavol Quittner, Karol Rečičár, Vladislav Rosa, František Rublík, Zita Sklenáriková, František Štulajter, Štefan Šujan, Valent Zaťko, *Archív Terminologickej komisie JSFM, 660 strán rukopisu*, Bratislava 2001
- [15] P. Quittner, *Variačné metódy*, rukopis, Bratislava FMFI UK 2005
- [16] *Matematická terminológia*, Slovenské pedagogické nakladateľstvo, Bratislava 1977
- [17] M. Ivanová-Šalingová, *Vreckový slovník cudzích slov*, Kniha-Spoločník, 1993
- [18] I. Búranský, *XML a webové služby.pdf*, Praha 2002
- [19] I. Kraval, *Základy XML.pdf*, Praha 2000
- [20] V. Sklenák, *XML a knihovny.pdf*, Praha 2001
- [21] <http://sk.wikipedia.org>, 15. 05. 2006
- [22] <http://en.wikipedia.org>, 15. 05. 2006
- [23] <http://www.w3.org/TR/xslt>, 15. 05. 2006
- [24] http://www.w3.org/Math/Software/mathml_software_cat_editors.html, 15. 05. 2006

- [25] <http://www.sccg.sk/~valentova>, 15. 05. 2006
- [26] <http://www.sccg.sk/~valentova/postup.php>, 15. 05. 2006
- [27] <http://www.sccg.sk/~valentova/xmleditor.php>, 15. 05. 2006
- [28] <http://validator.w3.org/>, 15. 05. 2006
- [29] <http://www.w3.org/1998/Math/MathML>, 15. 05. 2006
- [30] <http://www.forma.sk/onlines/slex/index.asp>, 15. 05. 2006
- [31] ISO/IEC 10646 – Univerzálna znaková množina

Príloha A

*Ak nerozumieš niektorému odbornému výrazu, nemaj strach!
Článok zostane aj bez neho úplne zrozumiteľný.
ZÁKON PÁNA COOPERA*

Definície pojmov a skratiek

- *Akronym*

Skratka pomenovania utvorená zo začiatočných hlások alebo slabík, prípadne ich kombináciou [17].

- *Cascading Style Sheets (CSS)*

Kaskádové štýly je všeobecné rozšírenie *HTML* o možnosti opisu vzhľadu textu základnými parametrami bežného *DTP*. Štýly umožňujú oddeliť štruktúru *HTML* alebo *XHTML* od vzhľadu [21].

- *Comma-separated values (CSV)*

Prehľadný dátový formát, ktorého polia sú oddelené čiarkou a hodnoty sú uzavreté v úvodzovkách [22].

- *Document Type Definition (DTD)*

Jedna z možností ako popísať štruktúru dokumentu. Určuje, ktoré elementy a atribúty budú k dispozícii a ako sa môžu navzájom kombinovať [20].

- *Frazeológia*

Súhrn ustálených spojení v danom jazyku alebo určitom jazykovom prejave, útvare [17].

- *International Standard Book Number (ISBN)*

Jedinečný identifikátor pre knihy [22].

- *HyperText Markup Language (HTML)*

Pôvodný návrh Tim Berners-Lee, potom odporúčanie W3C, posledná verzia 4.0 uverejnená v apríli 1998, aplikácia SGML pre zverejňovanie dokumentov na webe [18].

- *Lingvistika*

Veda o jazyku a jazykovej komunikácii; jazykoveda [17].

- *Namespaces in XML.*

Odporúčanie W3C uverejnené v januári 1999. Definuje spôsob pomenovania elementov XML dokumentu tak, aby nevznikali konflikty – nejednoznačnosti [18].

- *Normalizácia*

Zjednocovanie, zjednotenie predpisov, rozmerov, znakov, typov súčiastok, výrobkov; úprava, ustálenie smerníc, postupov pri určitej činnosti [17].

- *Optical character recognition (OCR)*

Softvér vytvorený na prevod obrázkov s textom (obvykle naskenovaných) do editovateľného textu [22].

- *Object-oriented programming (OOP)*

Objektovo orientované programovanie.

- *Standard Generalized Markup Language (SGML)*

Štandard ISO (International Organization for Standardization) s číselným označením 8879 z roku 1986 pre výmenu údajov medzi rôznymi počítačmi [18].

- *Tag*

Značka

- *Unicode*

Medzinárodný štandard, ktorého cieľom je definovať jedinečný kód zvaný code point pre každú z grafém používaných na zápis ľudského jazyka [22].

- *World Wide Web Consortium (W3C)*

Medzinárodné združenie organizácií pre tvorbu a správu štandardov pre World Wide Web. Zverejňuje ich na adrese <http://www.w3.org/TR> [18].

- *World Wide Web (WWW)*

Anglický názov pre časť celosvetovej počítačovej siete – internetu [21].

- *eXtensible Markup Language (XML)*

Odporúčanie W3C uverejnené vo februári 1998, druhá edícia uverejnená v októbri 2000, zjednodušenie SGML [18].

- *Extensible Stylesheet Language (XSL)*

Jazyk popisujúci transformáciu XML súborov a dát do ľubovoľného iného formátu [21].

- *XSL Transformations (XSLT)*

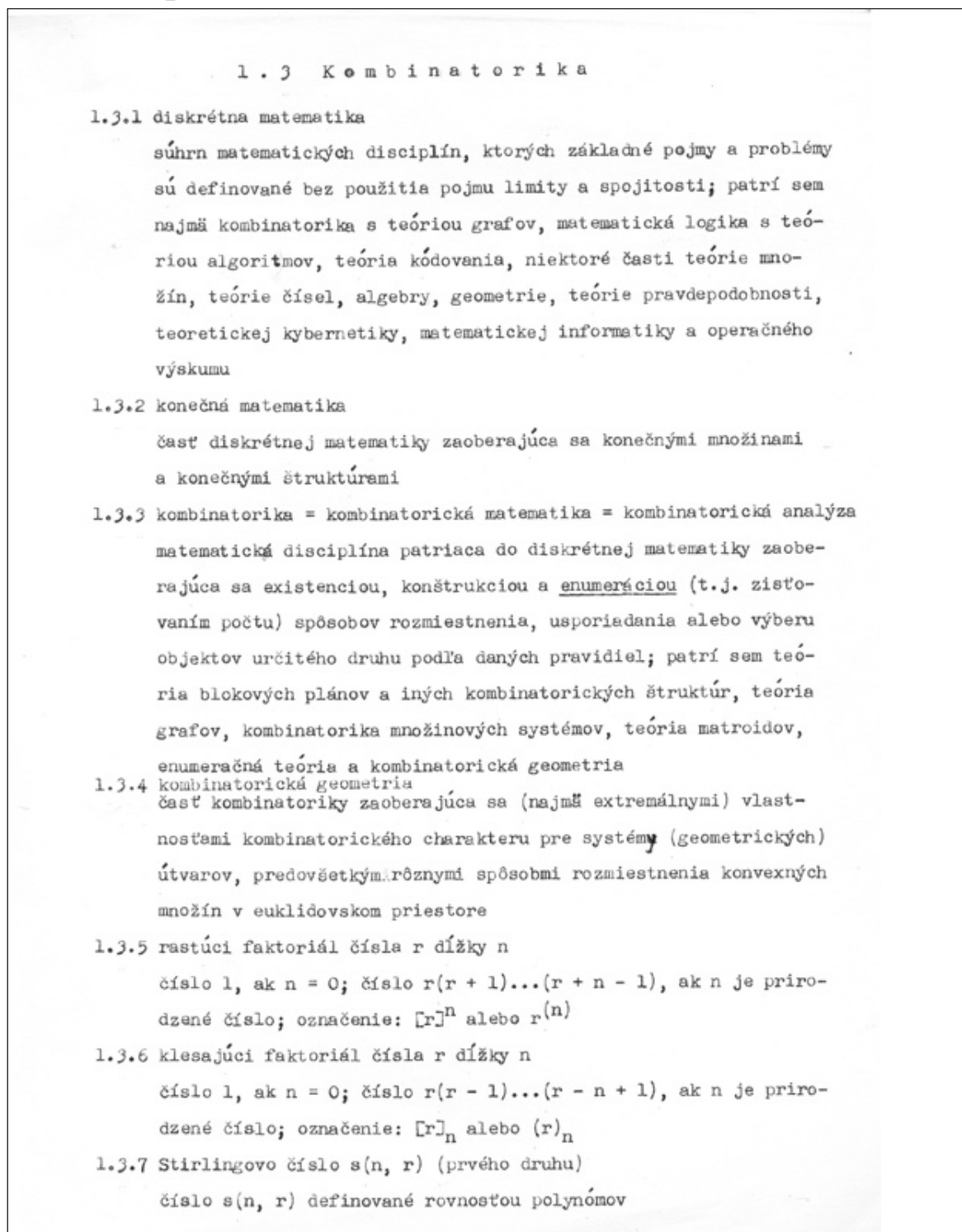
Jazyk na transformáciu XML dokumentu do iného XML dokumentu [23].

- *What You See Is What You Get (WYSIWYG)*

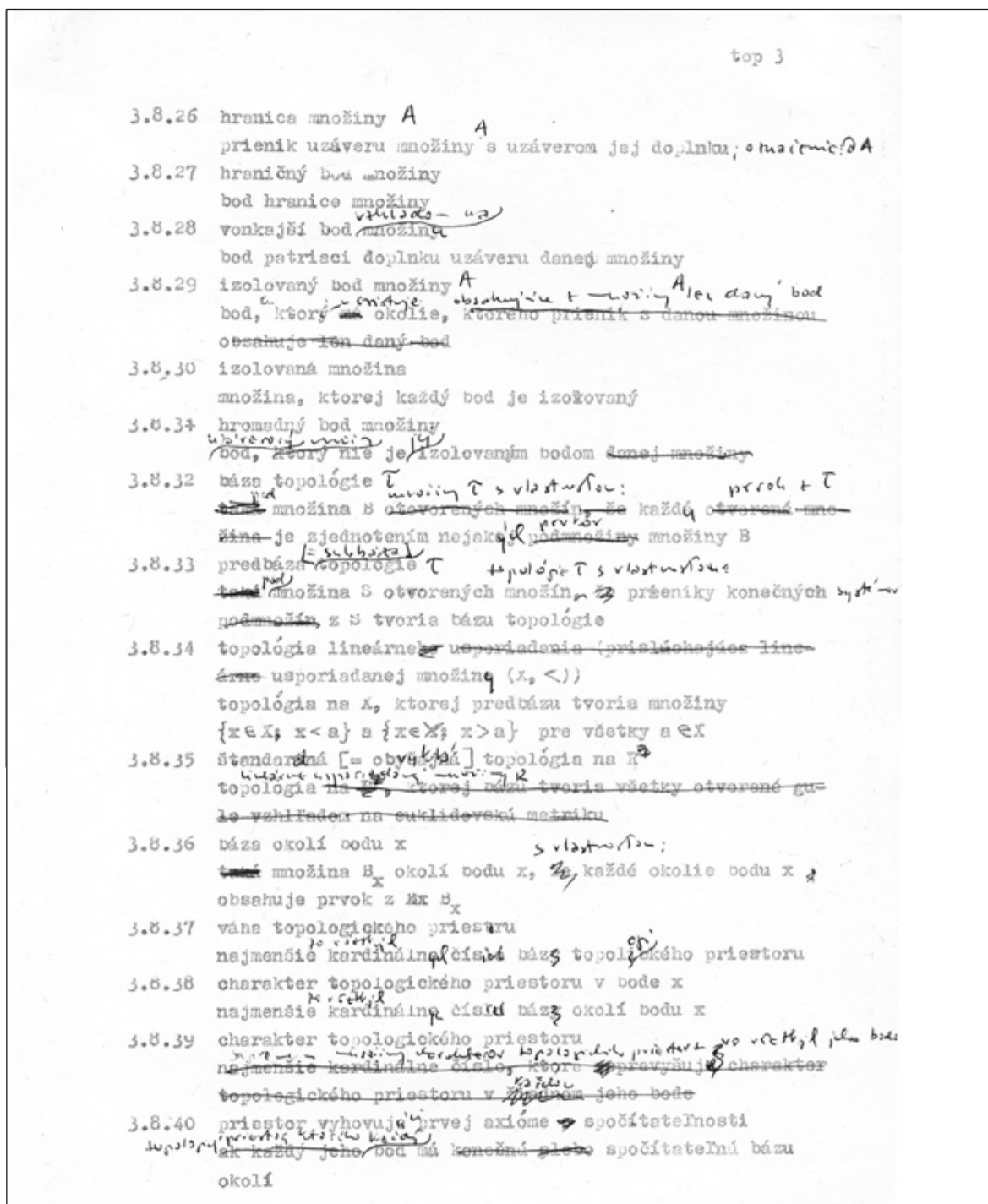
Princíp verného prenosu vizuálnej informácie, resp. informácie modelovanej na počítači do reality tak, že zodpovedá presne modelovanému obrazu s čo najmenším, resp. nebadateľným skreslením [21].

Príloha B

Obrazová príloha



Obrázok B.1:
Kvalitná strana.



Obrázok B.2:

Strana, na ktorej OCR zlyháva.

Príloha C

DVD – Obsah

Internetova_stranka	zdrojový kód stránky
Materialy – 660_stranova_kniha	obrázky
Rozpoznane_obrazky	<i>DOC, TXT, XML</i> súbory
Terminologicke_listy	obrázky
Ertl, Ferko – Slovník_pojmov.doc	termíny
Slovník.doc	termíny
SlovníkPG13SK_070305.doc	termíny
Setup	inštalačný súbor programu
Termíny	17 <i>XML</i> súborov, 17 <i>CSV</i> súborov, 3 <i>XSL</i> súbory
Text	<i>PDF</i> a <i>PPS</i> súbory
XML_Editor	zdrojový kód programu