



## POSUDEK VEDOUCÍHO DIPLOMOVÉ PRÁCE

**Jméno studenta:** Šárka Křížková

**Název práce:** Matematické metody pro získávání a zpracování informací

**Autor posudku:** doc. RNDr. Tatiana Gavalcová, CSc.

**Cíl práce:** Cílem diplomové práce je poskytnutí vhledu do problematiky získávání a zpracování informací (Information Retrieval, IR) se zaměřením na matematické metody používané v IR. Dále aplikovat vybrané matematické metody včetně jejich využití při detekci plagiátorství.

Povinná kritéria hodnocení práce	Stupeň hodnocení (známka)			
	1	2	3	4
Práce svým zaměřením odpovídá studovanému oboru	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Vymezení cíle a jeho naplnění	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Zpracování teoretických aspektů tématu	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Zpracování praktických aspektů tématu	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Adekvátnost použitých metod, způsob jejich použití	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Hloubka a správnost provedené analýzy	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Práce s literaturou	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Logická stavba a členění práce	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Jazyková a terminologická úroveň	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Formální úprava a náležitosti práce	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Vlastní přínos studenta	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Využitelnost výsledků práce v teorii (v praxi)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**Dílčí připomínky a náměty:**

V dalším textu.

**Celkové posouzení práce a zdůvodnění výsledné známky:**

Autorka strukturou i obsahem dodržuje okruh problematiky diplomové práce vymezený v jejím zadání. V úvodní části se podrobně věnuje popisu Information Retrieval: tímto termínem, ve velké míře používaném v anglickém jazyce, se na základě vymezení IR jako oboru vědeckého zkoumání rozumí způsoby vyhledávání a následného zpracování

informací. Pro účely oboru se předpokládá, že informace jsou uloženy v strukturovaných databázích a existují vhodně vyjádřené žádosti nebo dotazy uživatele o vyhledání právě takových informací, které jsou relevantní k jeho potřebám. Práce se věnuje konceptům v této oblasti, důležitým složkám problému zpracování informace, typologiím vyhledávacích požadavků a návazně na ně podává také časový, koncepční a částečně také technologický vývoj IR. Tato disciplína disponuje souborem nástrojů pro řešení problému podle odpovídajících modelů, ve kterých se určitým způsobem řeší také kvalita jeho výstupu. Autorka podává systematický přehled nejčastěji používaných modelů s jeho hlavním zástupcem - konceptuálním modelem.

Druhá část se zabývá matematickými modely jako nástroji v IR: pravděpodobnostním modelem, booleovským modelem a modely založenými na algebraických relačních strukturách nebo strukturách s definovanými operacemi (grafy, svazy, grupy, lineární prostory). Věnuje se principům využitelnosti modelů pro základní problém. V této části nejsou vyčerpány existující metody, které plynou z nabídky jiných matematických modelů; už například uplatnění samotné teorie svazů je úvodem do nové, několik desetiletí samostatně existující disciplíny označované jako Formal Concept Analysis. Další rozšíření by šlo ale do oblastí značně rozšiřujících zadané téma.

Třetí část práce se věnuje významné aplikaci metod IR, a to pro identifikaci plagiátorství. Uvádí koncepční principy detekce plagiátorství, jak je používají vybrané programy (např. Big Brother, Turnitin, Theses, Odevzdej a další).

Autorka prokázala schopnost nastudovat a zpracovat problematiku IR, kombinujícího problém získávání a zpracování informací pomocí relevantních nástrojů na bázi matematických metod, do uceleného odborného textu. Potřebnému studiu se autorka věnovala nad rámec odpovídajících předmětů v sylabech vysokoškolského studia. Získala vhled do charakteru úloh, do jejich aktuálního významu pro aplikace i do jejich rozdílné komplexity; témata lze v dalším rozvíjet. Považuji proto cíl diplomové práce za splněný. Postupovala přitom samostatně a systematicky. Zpracovala rozsáhlé množství odborné literatury, většinou v anglickém jazyce, samostatně dohledala relevantní tituly. Nepochybně už jen zachycené zdroje obsahují množství dalších témat nebo podrobností k problematice, zajímavých pro další zkoumání. Taková široká báze byla náročná na vlastní výběr zpracování, navíc byla nucena vyhledávat termíny ekvivalentní k odborným termínům v anglickém jazyce. Práce je napsaná obsahově i formálně přehledně a splňuje metodické pokyny i požadavky na diplomovou práci v tomto oboru.

**Práci doporučuji k obhajobě.**

**Otázky pro diskusi:**

Co bylo při přípravě diplomové práce pro vás největším přínosem z hlediska vlastního studia? Hodláte se problematikou IR dále nějakým způsobem zabývat?

**Navržená výsledná známka: výborně**

**V Hradci Králové, dne 12.5.2015**

---

**podpis**