

## Posudek oponenta diplomové práce

**Student:** Lengál Tomáš, Bc.

**Téma:** Segmentace webových stránek s využitím shlukování (id 19293)

**Oponent:** Bartík Vladimír, Ing., Ph.D., UIFS FIT VUT

- 1. Náročnost zadání** **průměrně obtížné zadání**  
Cílem diplomové práce bylo prostudovat metodu segmentace webových stránek založenou na shlukování a implementovat ji do nástroje FITLayout vyvíjeného na FIT. Metoda samotná sice není příliš složitá na pochopení, ale její implementace do FITLayoutu nebyla jednoduchým úkolem.
- 2. Splnění požadavků zadání** **zadání splněno**  
Všechny body zadání byly splněny.
- 3. Rozsah technické zprávy** **splňuje pouze minimální požadavky**  
Rozsah technické zprávy splňuje předepsaný požadavek na minimum 50 normostran.
- 4. Prezentací úroveň předložené práce** **80 b. (B)**  
Rozvržení práce do kapitol je v pořádku, stejně jako rozsahy jednotlivých kapitol. Po obsahové stránce bych asi ocenil, kdyby některé definice uvedené v kapitole 3 byly více komentovány nebo demonstrovány na příkladech - v těchto pasážích byla pochopitelnost textu o něco náročnější, v ostatních částech technické zprávy již byl text většinou srozumitelný.
- 5. Formální úprava technické zprávy** **62 b. (D)**  
Jazyková úroveň textu je podprůměrná, s řadu jazykových chyb a překlepů, ojediněle se zde vyskytují i větší hrubky. Po typografické stránce je technická zpráva až na pár drobností v pořádku.
- 6. Práce s literaturou** **90 b. (A)**  
Převzaté znalosti jsou v technické zprávě řádně odlišeny od vlastních výsledků studenta. Literatura je v textu práce řádně odkazována v souladu s příslušnými normami.
- 7. Realizační výstup** **90 b. (A)**  
Realizační výstup zahrnující implementovanou segmentační metodu je plně funkční a byl zahrnut do nástroje FITLayout. Byl také otestován na několika různých typech webových stránek. Oproti existujícímu známému algoritmu VIPS vyniká především rychlostí.
- 8. Využitelnost výsledků**  
Výsledná implementace bude mít přínos především pro další výzkum v oblasti extrakce informace a dolování z webových stránek na FIT.
- 9. Otázky k obhajobě**
  - Existuje nějaký způsob, jak lze uživateli doporučit vhodnou hodnotu parametru CT (shlukový práh)?
  - Jaké jsou výsledky implementované segmentační metody pro stránky neobsahující hlavní článek (např. osobní stránka zaměstnance fakulty)?
- 10. Souhrnné hodnocení** **80 b. velmi dobře (B)**  
Slabší stránkou této diplomové práce je technická zpráva, která má po formální stránce řadu nedostatků, vše je však vyváženo kvalitním realizačním výstupem. Proto navrhuji mírně nadprůměrné hodnocení stupněm B (velmi dobře).

Prohlášení: Uděluji VUT v Brně souhlas ke zveřejnění tohoto posudku v listinné i elektronické formě.

V Brně dne: 8. června 2017

.....  
podpis