

## Posudek oponenta bakalářské práce

**Student:** Panov Sergey  
**Téma:** Indexování a prohledávání sémanticky anotovaných textů (id 20056)  
**Oponent:** Dytrych Jaroslav, Ing., UPGM FIT VUT

- 1. Náročnost zadání** **obtížnější zadání**  
Zadání hodnotím jako obtížnější, protože bylo nutné se seznámit s větším množstvím nástrojů pro práci s velkými daty, detailně prostudovat existující indexační nástroj a jeho vnitřní struktury a provést v něm řadu změn. Současně bylo třeba navrhnout nového klienta a nový postup získávání dat ze serverů.
- 2. Splnění požadavků zadání** **zadání splněno**
- 3. Rozsah technické zprávy** **přesahuje obvyklé rozmezí**  
Práce má více než 60 normostran textu. Dle mého názoru je rozsah adekvátní řešené problematice, i když některé části teorie (např. v počátku 4. kapitoly) jsou dle mého názoru zbytečné či zbytečně podrobné.
- 4. Prezentací úroveň předložené práce** **89 b. (B)**  
Technická zpráva má logickou strukturu a jednotlivé kapitoly na sebe navazují. V rozsáhlejší teorii jsou vysvětleny všechny důležité pojmy a všechna rozhodnutí jsou dobře odůvodněna. Místa jsou popisy dle mého názoru zbytečně detailní, ale přehledné a srozumitelné. Jedinou slabinou je absence detailnějšího popisu nevýhod původního řešení, která vede k méně jasné motivaci pro některé změny.
- 5. Formální úprava technické zprávy** **99 b. (A)**  
Práce je jazykově čistá a byly dodrženy typografické zásady a doporučení.
- 6. Práce s literaturou** **92 b. (A)**  
Student čerpal převážně z vědeckých článků. Výběr literatury je relevantní pro dané téma, i když k některým popisovaným tématům existují i aktuálnější články.
- 7. Realizační výstup** **85 b. (B)**  
Programové řešení je plně funkční a připraveno na budoucí rozšíření o postfiltry. Funkčnost je ověřena jak provedenými testy na malé podmnožině dostupných datových sad, tak i reálným nasazením ve výzkumu. Dokumentační komentáře v kódu jsou však nedokončené (chybí popisy parametrů a návratových hodnot).
- 8. Využitelnost výsledků**  
Výsledky práce jsou již využívány v dalším výzkumu ve Výzkumné skupině znalostních technologií.
- 9. Otázky k obhajobě**
  - Proč jste v rámci práce neprovedl testy s celou datovou sadou Wikipedie či Common Crawl?
- 10. Souhrnné hodnocení** **89 b. velmi dobře (B)**  
Práce byla poměrně obtížná a časově náročná, protože bylo nutné pracovat s celou řadou nástrojů, nastudovat větší množství teorie, prostudovat vnitřní struktury složitého programu a navrhnout a implementovat řadu úprav i novou klientskou aplikaci. Vytvořené řešení bylo řádně otestováno, je plně funkční a bude dále využito.

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

V Brně dne: 25. května 2017

.....  
podpis