

Hodnocení vedoucího diplomové práce

Autor práce: **Petr Matějovič**

Název práce: **Rozšiřování dotazu pomocí sémantických prostorů**

Cílem předkládané práce bylo navrhnout, implementovat a testovat různé způsoby rozšiřování dotazu ve vyhledávacím systému tak, aby uživatel dostal lepší odpověď na dotaz původní. K rozšiřování dotazu měl student využít sémantické prostory (umožňují nalézt slova vyskytující se ve stejných kontextech).

Student k zadání přistoupil se zájmem, aktivně zkoumal a se mnou konzultoval dříve jemu neznámé principy a techniky související se zpracováním přirozeného jazyka, vyhledávacími systémy, vyhodnocením jejich kvality a sémantickými prostory.

Výborným výstupem studentovy práce jsou webová i konzolová aplikace umožňující pohodlně testovat několik studentem navržených, zajímavých a velmi různě nastavitelných způsobů rozšiřování dotazu na k tomuto účelu dodané standardní datové sadě. Testovat lze, velmi vhodně, i po dávkách. Výsledky jsou přehledně prezentovány, logovány a je možné je dále analyzovat. Mj. oceňuji přínosnou možnost zobrazení nastaveným systémem nenalezených relevantních článků k dotazům testovacích dat.

Již jen dobrým studentovým výstupem je pak testování způsobů rozšiřování dotazu. Student správně nevstřebal, a to hlavně z časových důvodů (mnoho času věnoval tvorbě aplikace), všechny souvislosti týkající vyhodnocení vyhledávacího systému. Proto testování, jeho popis v textu DP a z toho vyplývající závěry v DP nejsou zcela v pořádku.

Konkrétně student v práci porovnává systémy hlavně pomocí méně vhodných metrik, kterými jsou Přesnost a Úplnost, na nevhodně a neodůvodněně zvoleném počtu (200) vyhledávacím systémem vrácených dokumentů. V závěru pak student na základě kladných výsledků v těchto metrikách píše, že se „podařilo do jisté míry vylepšit vyhledávací systém“. S tím nesouhlasím. Souhlasil bych s bohužel naopak nezmíněným a zajímavým výsledkem: Podařilo se vytvořit systém, který sice uživateli nenabídne více relevantních článků mezi prvními vyhledanými dokumenty, ale takový, který nabídne uživateli více relevantních článků v prvních 200 vyhledaných. V kapitole testování dále chybí např. výpis všech testovaných parametrů, diskuse nad použitými mírami pro vyhodnocování nebo např. i popisky parametrů některých tabulek.

Na dokumentaci k DP se mi celkově líbí věcnost, množství názorných obrázků, tabulek a grafů. Mínusem dokumentace je, kromě již zmíněných nedostatků v kapitolách testování a závěr, větší množství faktických chyb a překlepů.

Ve shrnutí: Student splnil zadání. Navržené aplikační řešení je výborné – praktické, dále rozšiřitelné a nadstandardní. Teoretická i programátorská část dokumentace jsou velmi dobré. Testování a popis dosažených výsledků jsou dobré. Také vzhledem k obtížnosti zadání (mnoho souvisejících věcí, velká data, množství práce) a použitelnosti dodané aplikace pro další potenciální výzkum navrhuji hodnocení známkou lepší

Velmi dobře.

V Plzni dne 29.8. 2013

Ing. Lubomír Krčmář

**SOUHLASÍ
S ORIGINÁLEM**



Západočeská univerzita v Plzni
Fakulta aplikovaných věd
katedra informatiky a výpočetní techniky

②

