

# Posudek oponenta bakalářské práce

Autor práce: Lukáš Witz

Název práce: **Extrakce sociálních sítí ze zpravodajských textů**

Obsah práce:

Práce obsahuje popis principů základních metod NLP: segmentace, lemmatizace, stematizace, rozpoznávání jmenných entit a rozpoznávání polarity názoru. Následuje paragraf věnovaný zopakování základních pojmů z teorie grafů a popisu vybraných vizualizačních knihoven. Z textu na str.17 odvozují, že byly i testovány. Konkrétní výsledky testů nejsou uvedeny.

Druhá část práce obsahuje popis vytvořené aplikace SoNEx. Z poskytnutých dat, novinových článků, ve kterých byla označena vlastní jména, autor stvořil ohodnocený graf. Jeho hrany vyjadřovaly současný výskyt jmen ve větách, odstavcích, článcích. Konkrétní způsob ohodnocování jsem z textu nevyčetl. Bylo by užitečné dozvědět se také souhrnné údaje o vzniklém grafu. Autor mohl na úkor první části věnovat více místa popisu vlastního řešení. Graf byl pak použit k vyhledávání podgrafů ze zadaného uzlu prostřednictvím filtrovaného prohledávání do šířky a tím zjištěny entity mající souvislost s výchozím uzlem.

Kvalita řešení a dosažených výsledků:

Autor prokázal, že je schopen zvládnout řešený problém, seznámit se s relativně novou problematikou a zvládnout potřebné metody vyhledávání informací. Předložené výsledky vypadají věrohodně, způsob jejich ověření by zasluhoval sofistikovanější postup.

Formální úroveň

Je neobvyklé, aby se text abstraktů obsahově lišil. Práci by prospělo přesunutí části prostoru obecné druhé kapitoly ve prospěch stěžejní kapitoly třetí. Počet překlepů je přijatelný. Problémem je čitelnost grafů (např. obr.4.1).

Práce s literaturou:

Neshledal jsem nedostatky.

Splnění zadání:

**Zadání považuji za splněné ve všech bodech.**

Dotazy k práci:

Jaká možná rozšíření práce navrhuje? Proč myslíte, že graf se kterým pracujete (str.31) je souvislý?

Navrhuji hodnocení známkou výborně až velmi dobře a práci doporučuji k obhajobě.

V Plzni 16.5.2014

Prof. Ing. Karel Ježek, CSc.

