

# Úvaha o kľúčových slovách

Andrej Herich

Význam pojmu kľúčové slovo sa sprvoti javí ako zrejmý. Avšak pri pokuse formálne zachytiť tento význam, vzniknú problémy. Vágne povedané sú to slová, ktoré vystihujú podstatu dokumentu, článku, práce. Je ich možné využiť na zoskupovanie podobných dokumentov. Rovnako dobre je ale možné chápať kľúčové slová ako niečo, čo popisuje, čím sa dokument odlišuje od ostatných. Z tohto pohľadu by kľúčové slová mohli byť prirovnané k súradniciam dokumentu vo veľkom priestore všetkých dokumentov, kde tematicky blízke dokumenty sa nachádzajú blízko seba a tematicky odlišné dokumenty sa nachádzajú ďaleko od seba.

Jedným z problémov, ktoré sa týkajú kľúčových slov, je ich nejednoznačnosť. Na reprezentáciu kľúčového slova sa obvykle používa reťazec znakov zaznamenávajúci slovo v prirodzenom jazyku. To vedie k tomu, že pri viacerých dokumentoch môže byť jedno slovo použité vo viacerých významoch. Pri strojovom spracovaní sa budú takéto dokumenty javiť ako podobné, i keď obsahovo podobné nebudú. Lepším spôsobom reprezentácie by mohli byť významy slov prebrané zo slovníka alebo sémantickej siete (WordNet). Avšak automatické určenie významu slova je pomerne netriviálny proces a často sa stáva, že sa ani ľudia sa nevedia zhodnúť na tom, v ktorom význame je slovo v určitom kontexte použité. Myslím si, že ďalšie bádanie v tejto oblasti by mohlo priniesť dobré výsledky.