

Klicova slova si v podstate predstavuji jako neco skupinu urcitych slov, ktera doprovazi nejaky objekt – nejen text ci webovou stranku, ale napriklad i video. Klicova slova pak vyjadruji hlavni myslenku tohoto objektu – to, cim se tento objekt zabývá, o cem pojednava popr. s cim pracuje, co z nej vychazi. To je alespon to první, co se mi vybaví při pouziti vyrazu “Klicova slova.”

Potiz je vsak samozrejme v tom, ze klicova slova byvaji k temto objektum prirazovana, ve vetsine pripadu, samotnymi autory tohoto objektu. Autor si nemusí byt plne vedom ucelu techto klicovych slov popr. jejich vztahu k jinym objektum a dusledkum toho vztahu.

Jako jedno z moznych pouziti klicovych slov urcitych objektu je vzajemne porovnani techto objektu. Strojove porovnani celych objektu, napr. textu, slovo od slova, vetu od vety by nedavalo prilis smysl (nebo me smysl nenapada tak jasne jako u klicovych slov) a bylo jen tezce proveditelne - pocitac by nejdrive musel byt schopny vetam plne porozumet, aby mohl vyvodit zavery jine, nez napr. nejakou zakladni kategorizaci, kterou by vsak stejne musel nejspis provadet na zaklade nejakych jiz predem naucenych informaci. Pokud však jde o srovnani klicovych slov techto objektu, nabízí se mnohem jednodussi moznost porovnani techto objektu a vyvozovani jejich relevantnosti v tomto procesu. Dulezitou otazkou vsak stale zustava to, jak budou tato slova volena – jestli napr. jako nejcasteji se vyskytující slova, nebo slova takova, aby vystihovala podstatu objektu? Velmi casto se nemusí jednat o ta sama slova.

V praxi se pouzivala klicova slova casto ve webovem prostredi, ale z toho postupne seslo (alespon tedy v samotnych tazich html kodu, protoze to nebyl spolehlivy zdroj informaci). Dale si dokazi predstavit napriklad algoritmus, kdy se napriklad nejaky program pokusi rozdelit velkou sadu textu do ruznych kategorii (ktere sam vytvori), na zaklade klicovych slov. Hledal by tedy strukturu v neoznacenyh data – jako napr. ve strojovem uceni bez dohledu.