

Nome: Beatriz Ribeiro Borges

Matrícula: 12021BSI231

Organização e Recuperação da Informação

TP03 – Ex4

Discussão sobre os resultados das consultas feitas sem e com stemming.

Na consulta 1 “live me alive or alone” e na consulta 2 “enter the darkness of the twilight zone in my heart” apenas os primeiros lugares se mantiveram iguais. Contudo, houve uma mudança da similaridade: na consulta 1 teve um decréscimo com stemming e na 2 um pequeno aumento. “Beat it” e “I know theres something going on” continuam aparecendo do ranking da primeira consulta, mas em posições diferentes por conta da similaridade menor na versão com stemming.

Já a consulta 3 “every single time” continuou os mesmos documentos, a mesma ordem e aumentou um pouco a similaridade com stemming. É possível que isso tenha acontecido, pois as palavras já estão na sua menor forma.

A consulta 4 “sweet dreams are made of out of luck”, assim como a consulta 3, continuou com os mesmos documentos, a mesma ordem, mas os 2 primeiros lugares diminuíram um pouco a similaridade, enquanto os últimos 3 lugares aumentaram.

Por fim, a consulta 5 “i’ll work hard for money or for a pizza that i love” também permaneceu com os mesmos documentos e ranking iguais. O primeiro lugar teve um aumento da similaridade com stemming, provavelmente por ter “works” da música sendo transformado em “work” como na consulta. Os segundo e terceiro lugar teve um pequeno decréscimo. Já o quarto e quinto lugares, um aumento.

Com todas essas comparações, pode-se perceber que trabalhar com os radicais usando stemming (usando os radicais como base para cortar as palavras) faz com que haja mudanças, algumas vezes, significativas no ranking. Especialmente no grau de similaridade, mas também na ordem e seleção dos documentos em alguns casos.