

VSM with TF-IDF Weighting Still Has a Problem!

Query = “news about presidential campaign”

d1	... news about ...	$f(q,d1)=2.5$
d2	... news about organic food campaign...	$f(q,d2)=5.6$
d3	... news of presidential campaign ...	$f(q,d3)=7.1$
d4	... news of presidential campaign presidential candidate ...	$f(q,d4)=9.6$
d5	... news of organic food campaign... campaign...campaign...campaign...	$f(q,d5)=13.9?$