

Written Assignment 4
CS 510: Information Retrieval on the Internet
Due Tuesday May 11, 2010
8 points

Question A: (3 points) Answer Exercise 9.3 in the text. Show the *original* query vector as well as the modified query.

Question B: (4 points) Order the following four terms by their Normalized Google Distance (NGD) to the term *porcupine*. Indicate what you are using for N in your calculations.

quill
meatballs
hedgehog
cactis

Repeat the process using the Normalized Bing Distance (or for any other search engine that provides counts of matching documents).

Question C: (1 point) Consider the Heatmap for Boeing Corp. at <http://www.textmap.com/company/boeing.htm>

Why is there “heat” in the center of the country? Why isn’t there much “heat” around Chicago (which is Boeing’s headquarters).