Assignment 6

Searching
Document classification

Due Monday, November 7

1. Do question #11 in chapter 4 of the draft textbook (p. 152).

2. Do problem 4 from: https://sites.google.com/site/languageandcomputersfall2010/assignments/homework-2, but replace Austin, Texas with Bloomington, Indiana

3. From the same website, do problem 5.

4. Do question #1 from chapter 5 of the draft textbook (p. 190–191).

5. Select two articles in the Indiana Daily Student (http://www.idsnews.com/) written by different authors.
   (a) Give the title, author’s name, and date of each article.
   (b) What are some prominent stylistic differences between the two authors? Describe at least 3 differences.
   (c) How could these differences be detected automatically?

6. Write a Python program that takes two strings and says what their shared "prefix" is, if any. For example, bring and behold have a shared prefix of b
   • Let’s call the two strings a and b and assume that the first string is always shorter—i.e., len(a) <= len(b)
   • I (will) have some code available on the website to get you started.