Asst2

invertedIndex

Introduction: In this assignment we practiced using the file system API (as well as pointers in different data structures). In particular you will be creating, opening, reading, writing, and deleting files. Your task is to write an indexing program, called an indexer. Given a set of files, an indexer will parse the files and create an inverted index, which maps each token found in the files to the subset of files that contain that token. In your indexer, you will also maintain the frequency with which each token appears in each file. The indexer should tokenize the files and produce an inverted index of how many times the word occurred in each file, sorted by word. Your output should be in the following format: <?xml version="1.0" encoding="UTF-8"?> count0 count1 count2 count3 count4