Kevin Nguyen, Jacqueline Wong

Team #4

CS 6301 - Software Analysis and Comprehension

**Assignment 3**

**Apache Lucene**

* **# of JIRAs**: 7066
* Title Corpus
  + Sentence Average: 1
  + Sentence Median: 1
  + Term Average: 8
  + Term Median: 7
  + Noun Average: 4
  + Noun Median: 5
  + Verb Average: 0
  + Verb Median: 1
  + Adjective Average: 0
  + Adjective Median: 1
  + Adverb Average: 0
  + Adverb Median: 0
  + Top 10 Most Frequent Nouns (Greatest to Least): Index, Query, Field, test, term, Filter, -rrb-, file, patch, search
  + Top 10 Most Frequent Verbs (Greatest to Least): add, remove, make, fail, improve, allow, create, set, build, merge
  + Top 10 Most Frequent Adjectives (Greatest to Least): token, simple, spatial, wrong, open, multiple, incorrect, equal, small, false
  + Top 10 Most Frequent Adverbs (Greatest to Least): not, correctly, incorrectly, properly, close, fast, backwards, back, more, directly
* Description Corpus
  + Sentence Average: 4
  + Sentence Median: 3
  + Term Average: 82
  + Term Median: 48
  + Noun Average: 37
  + Noun Median:22
  + Verb Average: 6
  + Verb Median: 5
  + Adjective Average: 11
  + Adjective Median: 6
  + Adverb Average: 1
  + Adverb Median: 1
  + Top 10 Most Frequent Nouns (Greatest to Least): -rrb-, test, text, junit, index, -lrb-, code, term, field, file, query
  + Top 10 Most Frequent Verbs (Greatest to Least): be, add, do, create, make, fail, remove, build, merge, think
  + Top 10 Most Frequent Adjectives (Greatest to Least): token, null, true, false, current, different, open, single, simple, new
  + Top 10 Most Frequent Adverbs (Greatest to Least):not, currently, actually, back, simply, here, probably, directly, faster, close
* Title + Description Corpus
  + Sentence Average: 4
  + Sentence Median: 3
  + Term Average: 88
  + Term Median: 50
  + Noun Average: 41
  + Noun Median: 27
  + Verb Average: 8
  + Verb Median: 6
  + Adjective Average: 12
  + Adjective Median: 7
  + Adverb Average: 1
  + Adverb Median: 1
  + Top 10 Most Frequent Nouns (Greatest to Least): -rrb-, test, text, index, junit, term, field, -lrb-, code, file
  + Top 10 Most Frequent Verbs (Greatest to Least): be, add, do, create,fail, remove, make, build, merge, make
  + Top 10 Most Frequent Adjectives (Greatest to Least): token, null, true, false, current, open, different, simple, single, new
  + Top 10 Most Frequent Adverbs (Greatest to Least): not, currently, actually, back, simply, here, probably, directly, correctly, faster

**Apache Ofbiz**

* **# of JIRAs**: 6917
* Title Corpus
  + Sentence Average: 1
  + Sentence Median: 1
  + Term Average: 9
  + Term Median: 7
  + Noun Average: 5
  + Noun Median: 5
  + Verb Average: 0
  + Verb Median: 1
  + Adjective Average: 0
  + Adjective Median: 0
  + Adverb Average:0
  + Adverb Median: 0
  + Top 10 Most Frequent Nouns (Greatest to Least): error, order, screen, product, work, item, service, form, field, page
  + Top 10 Most Frequent Verbs (Greatest to Least): add, create, find, remove, show, render, update, miss, improve, generate
  + Top 10 Most Frequent Adjectives (Greatest to Least):
  + Top 10 Most Frequent Adverbs (Greatest to Least): not, correctly, properly, bakc, incorrectly, twice, automatically, better, anymore, manually
* Description Corpus
  + Sentence Average: 4
  + Sentence Median: 3
  + Term Average: 61
  + Term Median: 33
  + Noun Average: 36
  + Noun Median:18
  + Verb Average: 7
  + Verb Median:4
  + Adjective Average: 9
  + Adjective Median:3
  + Adverb Average: 0
  + Adverb Median: 0
  + Top 10 Most Frequent Nouns (Greatest to Least): order, file, error, service, field, product, screen, type, component, name
  + Top 10 Most Frequent Verbs (Greatest to Least): create, be, add, find, do, id, follow, read, set, get
  + Top 10 Most Frequent Adjectives (Greatest to Least): main, null, true, select, unknown, current, false, different, complete, key
  + Top 10 Most Frequent Adverbs (Greatest to Least): not, back, currently, probably, actually, correctly, directly, automatically, here, only
* Title + Description Corpus
  + Sentence Average: 4
  + Sentence Median: 3
  + Term Average: 69
  + Term Median: 37
  + Noun Average: 41
  + Noun Median:24
  + Verb Average: 8
  + Verb Median: 5
  + Adjective Average: 10
  + Adjective Median: 4
  + Adverb Average: 1
  + Adverb Median: 0
  + Top 10 Most Frequent Nouns (Greatest to Least): order, file, error, service, product, screen, field, type, component, item
  + Top 10 Most Frequent Verbs (Greatest to Least): create, add, be find, id, do, follow, read, set, get
  + Top 10 Most Frequent Adjectives (Greatest to Least): main, null, true, select, unknown, current, false, different, complete, key
  + Top 10 Most Frequent Adverbs (Greatest to Least): not, back, correctly, probably, currently, actually, automatically, directly, properly, here