CorpusReader

String[] ignoreExtensions; ArrayList<CorpusFiles> files;

readDir(String path)
ArrayList<CorpusFile> getFiles()
CorpusFile getNextFile()

CorpusFile

String path; String extension;

String getPath(); String getExtension();

DocumentProscessor

ArrayList<Documents> documents

readDir(String path)
ArrayList<CorpusFile> getFiles()
CorpusFile getNextFile()

Document

static String tmpPath; static Int id;

Int docld;

Int docStartLine;

String filePath;

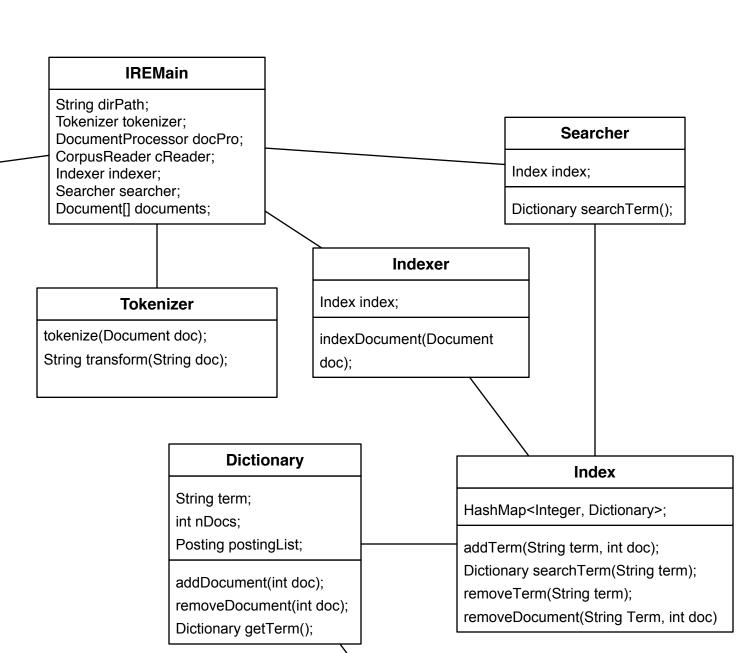
String documentPath;

Int getDocStartLine();

Int getDocId();

String getDocumentPath();

String getFilePath();



Postings

ArrayList<Integer> docs;

addPosting(int doc);

int postingListSize();

 $remove Posting (int \ doc);\\$

ArrayList getAllPostings();