Zadatak Koristeći Apache Lucene biblioteku indeksirati kolekciju tekstualnih fajlova i isprobati opcije za rangiranje rezultata primenom mera sličnosti (*similarity*).

1. Kao osnovu iskoristiti projekat i kolekciju tekstualnih fajlova iz prve laboratorijske vežbe (koristiti samo originalne kolekciju od 4 fajla **bez druge kolekcije** dobijene deljenjem fajlova iz originalne kolekcije). **Obavezno je da svi studenti imaju različite kolekcije fajlova sa kojima rade.**
2. Kreirati jedan indeks nad tim fajlovima kao u prvoj laboratorijskoj vežbi koristeći podrazumevanu klasu za meru sličnosti - **BM25Similarity**.
3. Dopuniti prvu laboratorijsku vežbu tako da se za svaki dokument koji je rezultat pretrage prikazuje *score* koji se može dobiti iz atributa *org.apache.lucene.search.ScoreDoc.score* i objašnjenje za *score* upotrebom metode *Explanation explain(Query query, int doc)* iz klase *org.apache.lucene.search.IndexSearcher*.
4. Osmisliti Bulovski upit od najmanje 2 termina koji vraća rezultat kad se izvrši nad poljem naslov i nad poljem sadržaj. Npr. nad kolekcijom fajlova iz projekta *QueryTester* sa računskih vežbi moguće je izvršiti sledeće upite (važno je da postoji jedan dokument koji vraćaju i jedan i drugi upit, a moguće je da neki od njih vrati i više dokumenata):
    a. naslov:Anna naslov:Karenina
    b. sadrzaj:Anna sadrzaj:Karenina
5. Proučiti *score* ta dva upita, kako je dobijena ta vrednost kroz objekat *Explanation*, kako dužina polja (polje sadržaj je mnogo duže od polja naslov) utiče na *score* itd. Za upit od ta dva koji vraća manji *score* za dokument podesiti *boost* vrednost tako da oba upita (i nad sadržajem i nad naslovom) imaju isti *score* za taj dokument. Za prethodne primere upita zajednički dokument za oba slučaja je *Anna Karenina.txt*. Za upit po naslovu *score* je 1.0945207, a za upit po sadržaju *score* je 1.5218642.
6. **Ponoviti tačke 2., 3., 4. i 5.** tako da je prilikom indeksiranja i prilikom pretrage kao mera sličnosti postavljena klasa **ClassicSimilarity** (tj. kreirati novi indeks, ponovo izvršiti iste upite i podesiti novu vrednost za *boost*).
7. Predati preko CS-a **isključivo ZIP arhivu sa projektom** (**ne .rar, .7z**, itd.) Pre zipovanja projekta **obrisati JAR fajlove, .CLASS fajlove i kreirani indeks**. Treba **predati samo sopstveni Java kod**, **tekstualne fajlove** i **projektne fajlove iz razvojnog okruženja**). Ova ograničenja su važna da ukupna veličina arhive **ne bi prešla 1MB koliki je limit na CS-u**.