Συστήματα Ανάκτησης Πληροφοριών

Φάση 1 – Baseline – Μοντέλο Ανάκτησης Διανυσματικού Χώρου

Στυλιανή Δούκα – p3170042

Βασίλειος Μπάλλας – p170115

**Περιγραφή Υλοποίησης**

1. Για την αποθήκευση των κειμένων στο ευρετήριο χρησιμοποιήσαμε τα πεδία *.I, .T, .A, .W.* Πιο συγκεκριμένα, για κάθε κείμενο αποθηκεύσαμε το μοναδικό κωδικό (id) του: *.I*, τον τίτλο (title) του: *.Τ*, τον συγγραφέα(author) του: *.A* και το κείμενο (content) του: *.W*. Όλα τα παραπάνω αποθηκεύτηκαν σε *TextFields.* Για κάθε ερώτημα (*query*) αποθηκεύσαμε το μοναδικό κωδικό (id) του: *.I* και το περιεχόμενο (content) του: *.W*.
2. Για την επεξεργασία των κειμένων χρησιμοποιήθηκε ο *EnglishAnalyzer*,τοσο για τα κείμενα που αποθηκεύονται στο ευρετήριο, όσο και για τα ερωτήματα (*queries*) που θα χρησιμοποιηθούν για την αναζήτηση. Ως συνάρτηση ομοιότητας, μεταξύ κειμένων και ερωτημάτων, χρησιμοποιήθηκε η *ClassicSimilarity*.
3. 