

Σύστημα Ανακήτησης Πληροφοριών
HY-463
Πανεπιστήμιο Κρήτης

Α' φάση Πρότζεκτ

*Σταυρακαντωνάκη Ειρήνη
Βογιατζής Γιάννης*

Το πρώτο μέρος της εργασίας ήταν η κατασκευή ενός συστήματος ανάκτησης πληροφοριών χρησιμοποιώντας έτοιμα έγγραφα. Η φάση αυτή αποτελούνταν από δυο μέρη :

- Ευρετηρίαση
- Αποτίμηση Επερωτήσεων

Ευρετηρίαση

Η διαδικασία αυτή περιελάμβανε 6 βήματα που αφορούσαν

- Την ανάγνωση των αρχείων και αποθήκευση των περιεχόμενων σε δομές σχεδιασμένες από εμάς
- Stemming του λέξεων
- Δημιουργία των φακέλων που θα χρησιμοποιηθούν στην Αποτίμηση

Κλάσεις που χρησιμοποιήθηκαν:

1. Document
2. Indexer
3. Parser
4. Word
5. WordFrequency

Αποτίμηση Επερωτήσεων

Στο μέρος αυτό είχαμε να υλοποιήσουμε τα εξής

- Αποτίμηση ερωτήσεων βάσει του inverted index
- Αποτίμηση ερωτήσεων βάσει του Vector Space
- Αποτίμηση ερωτήσεων βάσει του Okapi BM25
- Snippet

Κλάσεις που χρησιμοποιήθηκαν:

1. Query Evaluator
2. Query GUI
3. QueryResults
4. RetrievalModel
5. ScoreEntry
6. VectorSpaceModel
7. OKAPIBM25
8. Vocabulary
9. VocabularyEntry
10. NoneRetrievalModel
11. ScoreEntry