

Songs Search Engine

Ιωάννης Τσιρώνης 4908

Κωνσταντίνος Ανδρέου 4316

[GitHub](#)

README

Το Project είναι χωρισμένο σε δυο κύρια μέρη:

- **Django Application**
- **Lucene Application**

Το **Django Application** χωρίζεται σε:

- **Frontend**
- **Backend**

- Το **Frontend** διαχειρίζεται την επικοινωνία του χρήστη με την εφαρμογή και για την υλοποίηση του χρησιμοποιούμε Django Templates.

- Το **Backend** διαχειρίζεται τις επιλογές του χρήστη από το Frontend και είναι υπεύθυνο για την επικοινωνία με το Lucene Application(SearchEngine\services search.py). Επίσης, διαχειρίζεται την επεξεργασία των αποτελεσμάτων που δέχεται από το Lucene Application(SearchEngine\services process_output.py) για την παρουσίαση τους στον τελικό χρήστη.

Το **Lucene Application** έχει το μορφή ενός .jar αρχείου το οποίο καλείτε από το Backend με τα κατάλληλα ορίσματα για να γίνει η κάθε αναζήτηση. Αποτελείται από δυο αρχεία:

- CreateIndex.java για τη δημιουργία των indexes(Καλείται μια φορά από τον administrator και δεν ξανακαλείται).
- LuceneModule.java καλείται από το Backend κάθε φορά που δέχεται αίτημα για αναζήτηση στη βάση και επιστρέφει τα top 10 αποτελέσματα πίσω στο service που το κάλεσε.

Το αρχείο /DataPreprocessing.ipynb είναι υπεύθυνο για την προ-επεξεργασία των δεδομένων που πήραμε από το Kaggle.

Στο `data/example_songs.tsv` βρίσκονται 20 τυχαία τραγούδια από το dataset που χρησιμοποιήθηκε στην κανονική εφαρμογή.

Αυτή ήταν μια σύντομη περιγραφή της δομής του project και δεν περιγράφεται το πώς λειτουργούν όλα αυτά μιας και επρόκειτο να αλλάξουν στη διαδικασία της εργασίας.