

# README

This is a domain specific search retrieval system. It searches any dialogue from the book "The Time Machine " and will give you the page number.

Run the files in the following order:

- Tokenisation.py
- Inverted\_index.py
- Words.py
- tf-idf\_values.py
- scoring.py
- trail.py

or Running trail.py would be enough as all the json files already have data in them.

After running trail.py in terminal a statement saying "enter the dialogue:" will be printed and underneath which user need to enter the dialogue that he wants to search for. Then the results of top 10 will be shown.

All the documents are stored in the folder corpus.

All the json files are stored in the folder savers.

TEAM MEMBERS:

- 1.CHINTHAKINDI SIDHARTHA RAMAN (2018A7PS0283H)
- 2.POOJARI VENKATA SHREESH (2018A7PS0301H)
- 3.BEJANKI ADITYA (2018A7PS0282H)
- 4.DINESH REDDY NANDYALA (2018A7PS0431H)