
TEXT MINING AND RELATIONSHIP EXTRACTION TO INFORM POLICY RECOMMENDATIONS AND INCREASE IMPACT

Nationaal Rapporteur

PwC - Nationaal Rapporteur

Overview

Goals

- Read in all files and put into database
- Implement a spell check based on documents
- Implement TFIDF search algorithm
- Implement LSI search algorithm
- Create an app to visualise how searching works

Summary

Accomplished

- Extract text from files
- Add extracted data to 'dataframe'
- LSI Search
- Spell Check
- Search by Synonyms
- Return results by page, instead returning documents

To Be Added

- LDA Search
 - UI
 - Link topics with search terms
-