

Document Analysis Tool

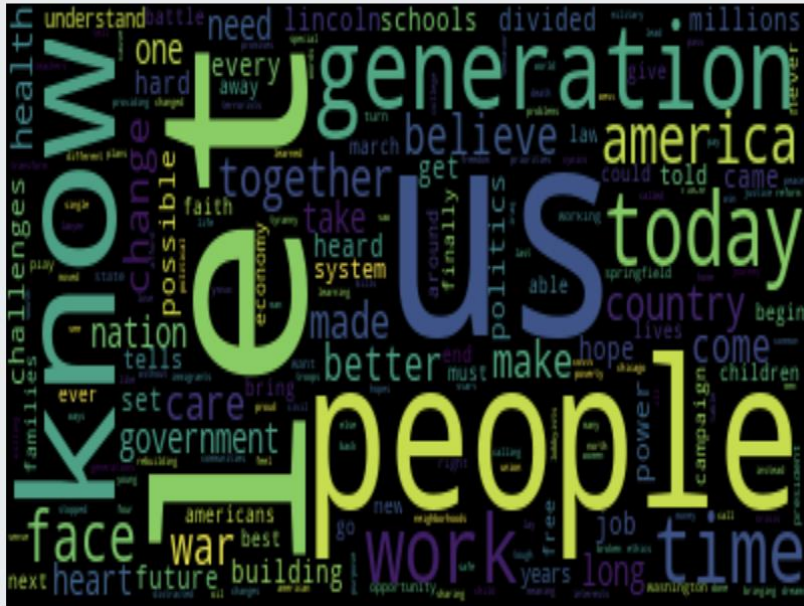
- This is a flask application for text folder analysis.

Sample result (built 5 models)

Analysis Result for document 1

1. Model1 – word frequency word cloud

word cloud



2. Mode2 - key words extraction by tf-idf algorithm

key words extraction

- party
- enough
- america
- know
- care
- saw
- time
- country
- say
- people

3. Mode3 - key phrases extraction by RAKE algorithm

key phrases extractionn

- actually two different levels offered
- saw four little girls die
- watched immigrants leave distant shores
- single mom whose two jobs
- new world comes new risks

4. Mode4 -topic related to document by LDA algorithm

topic model

- us
- america
- one
- time
- promise
- country
- american
- people
- know
- care

5. Model5 - document correlation matrix (input matrix is LDA model result)

	1	3	2	6	5	4
1	1.000000	-0.111111	-0.111111	-0.111111	-0.111111	-0.111111
3	-0.111111	1.000000	1.000000	1.000000	-0.111111	-0.111111
2	-0.111111	1.000000	1.000000	1.000000	-0.111111	-0.111111
6	-0.111111	1.000000	1.000000	1.000000	-0.111111	-0.111111
5	-0.111111	-0.111111	-0.111111	-0.111111	1.000000	-0.111111
4	-0.111111	-0.111111	-0.111111	-0.111111	-0.111111	1.000000

How to use:

- At upload file page
- Enter the document id, user wants to analyze
- Upload the whole document folder.

[Home](#) [Upload File](#)

NLP tool will analyze text folder and return the analytical result

document number you want to analyze:

upload your text folder here:

6 files

How to run:

cd to the **doc_analysis** folder

\$ python runserver.py