FY2020 Quick Search Analysis

GC Document Processing to JSON required before hand

```
import json
import pandas as pd
with open('NDAA_FY2020.json') as f:
  data = json.load(f)
```

Defining Keywords for analysis

Current keywords: transform, reform

Choosing stem words to catch the largest amount of variance

```
keywords = ['transform', 'reform']
key_dict = {}
```

```
for par in data['paragraphs']:
    for key in keywords:
        if key in par['par_raw_text_t']:
            if key in key_dict:
                key_dict[key].append(par)
            else:
                key_dict[key] = [par]
```

```
all_df = []
out_dict = []
for keyword in keywords:
    for item in key_dict[keyword]:
        entity = []
        for ent in item['entities']:
            for ent_s in item['entities'][ent]:
                entity.append(ent_s)
        out_dict.append({
            'keyword': keyword,
            'filename':item['filename'],
            'page_num':item['page_num_i'],
            'par_raw_text_t':item['par_raw_text_t'],
            'entities':entity,
        })
df = pd.DataFrame(out_dict)
df
```

keyword

filename

page_num

par_raw_text_t

entities

0

transform

NDAA_FY2020.pdf

110

H . R . 6395—111 (d)AGENCY PARTICIPATION.—Th…

[the Food and Drug Administration, the Nationa…

1

transform

NDAA_FY2020.pdf

353

H . R . 6395—354 and applicability of such tec…

[the Chief Information Office, the Defense Adv…

2

transform

NDAA_FY2020.pdf

551

H . R . 6395—552 ‘ ‘ (vi) transformation of …

[State, the Department of Defense, SEC, C the …

3

transform

NDAA_FY2020.pdf

657

H . R . 6395—658 (B) notify the appropriate …

[National Security Space Launch, the Space For…

4

transform

NDAA_FY2020.pdf

1065

H . R . 6395—1066 SEC . 4201 .RESEARCH , DEVEL…

[T&E, RAND ARROYO CENTER, MLRS, CHINOOK, SEC, …

5

reform

NDAA_FY2020.pdf

2

H . R . 6395—3 Sec . 212 .Disclosure requireme…

[the Joint Artificial Intelligence Center, the…

6

reform

NDAA_FY2020.pdf

10

H . R . 6395—11 Sec . 757 .Study on force mix …

[Department of Defense, Department of Veterans…

7

reform

NDAA_FY2020.pdf

11

H . R . 6395—12 Sec . 838 .Comptroller General…

[Congress, the Small Business Technology Trans…

8

reform

NDAA_FY2020.pdf

23

H . R . 6395—24 Sec . 2503 .Execution of proje…

[Department of Defense, Department of Defense …

9

reform

NDAA_FY2020.pdf

66

H . R . 6395—67 Sec . 212 .Disclosure requirem…

[the Joint Artificial Intelligence Center, the…

10

reform

NDAA_FY2020.pdf

96

H . R . 6395—97 Defense as the Secretary consi…

[the Joint Artificial Intelligence Center, D Q…

11

reform

NDAA_FY2020.pdf

334

H . R . 6395—335 (4)The effect of any such l…

[the Federal Government, nent, United States C…

12

reform

NDAA_FY2020.pdf

340

H . R . 6395—341 (3)The increase , as compar…

[Department of Defense, the Joint Artificial I…

13

reform

NDAA_FY2020.pdf

341

H . R . 6395—342 Sec . 832 .Extension of pilot…

[Department of Defense, the Small Business Tec…

14

reform

NDAA_FY2020.pdf

359

H . R . 6395—360 (5)SECRETARY CONCERNED.—The…

[the Department of Defense, SEC, United States…

15

reform

NDAA_FY2020.pdf

373

H . R . 6395—374 SEC . 838 .COMPTROLLER GENERA…

[the Defense Science Board Task Force, the Dep…

16

reform

NDAA_FY2020.pdf

412

H . R . 6395—413 an analysis of the process us…

[Congress, Subtitle BOther Department of Defen…

17

reform

NDAA_FY2020.pdf

413

H . R . 6395—414 of the Department of Defense …

[Congress, the Department of Defense, SUBMITTA…

18

reform

NDAA_FY2020.pdf

414

H . R . 6395—415 General of the United States …

[Department of Defense, the Department of Defe…

19

reform

NDAA_FY2020.pdf

543

H . R . 6395—544 ‘ ‘ (1) A staffing plan to …

[Congress, Hizballah, Islamic Revolutionary Gu…

20

reform

NDAA_FY2020.pdf

551

H . R . 6395—552 ‘ ‘ (vi) transformation of …

[State, the Department of Defense, SEC, C the …

21

reform

NDAA_FY2020.pdf

553

H . R . 6395—554 (D) an assessment of the ex…

[Congress, State, air force, the Committee on …

22

reform

NDAA_FY2020.pdf

568

H . R . 6395—569 (6) affirms the commitment …

[State, Congress, the Department of Defense, S…

23

reform

NDAA_FY2020.pdf

580

H . R . 6395—581 (C)1 member selected by agr…

[TRANSITIONAL PERIOD.The, the General Intellig…

24

reform

NDAA_FY2020.pdf

581

H . R . 6395—582 intelligence services and str…

[SEC, the Child Soldiers Prevention Act, Congr…

25

reform

NDAA_FY2020.pdf

587

H . R . 6395—588 of title 10 , United States C…

[Department of Defense, United Nations, SEC, t…

26

reform

NDAA_FY2020.pdf

589

H . R . 6395—590 involved in the illicit trade…

[State, SEC, Fed, FORM.The, Cabinet, the Sover…

27

reform

NDAA_FY2020.pdf

928

H . R . 6395—929 Subtitle B—Military Family Ho…

[Department of Defense, Guam Realignment, Neva…

28

reform

NDAA_FY2020.pdf

1112

H . R . 6395—1113 SEC . 4501 .OTHER AUTHORIZAT…

[PROC REPLACEMENT & MODERNIZATION, OFFICE OF T…

Writing simple table to file

Output File is in attached CSV at NDAA_FY2020_QuickSearch.csv

```
df.to_csv('NDAA_FY2020_QuickSearch.csv')
```

Further Analysis