ndaa_book

FY2020 Quick Search Analysis¶

GC Document Processing to JSON required before hand¶

In [74]:

```python
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

In [75]:

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

In [76]:

```python
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]
```

In [77]:

```python
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
```

Out[77]:

	keyword	filename	page_num	par_raw_text_t
0	transform	NDAA_FY2020.pdf	110	H . R . 6395—111 (d)AGENCY PARTICIPATION.—
1	transform	NDAA_FY2020.pdf	353	H . R . 6395—354 and applicability of such tec...
2	transform	NDAA_FY2020.pdf	551	H . R . 6395—552 ' ' (vi) transformation of ...
3	transform	NDAA_FY2020.pdf	657	H . R . 6395—658 (B) notify the appropriate ...
4	transform	NDAA_FY2020.pdf	1065	H . R . 6395—1066 SEC . 4201 .RESEARCH , DEVEL.
5	reform	NDAA_FY2020.pdf	2	H . R . 6395—3 Sec . 212 .Disclosure requireme...
6	reform	NDAA_FY2020.pdf	10	H . R . 6395—11 Sec . 757 .Study on force mix ...
7	reform	NDAA_FY2020.pdf	11	H . R . 6395—12 Sec . 838 .Comptroller General...
8	reform	NDAA_FY2020.pdf	23	H . R . 6395—24 Sec . 2503 .Execution of proje...
9	reform	NDAA_FY2020.pdf	66	H . R . 6395—67 Sec . 212 .Disclosure requirem...
10	reform	NDAA_FY2020.pdf	96	H . R . 6395—97 Defense as the Secretary consi...
11	reform	NDAA_FY2020.pdf	334	H . R . 6395—335 (4)The effect of any such l...
12	reform	NDAA_FY2020.pdf	340	H . R . 6395—341 (3)The increase , as compar...
13	reform	NDAA_FY2020.pdf	341	H . R . 6395—342 Sec . 832 .Extension of pilot...
14	reform	NDAA_FY2020.pdf	359	H . R . 6395—360 (5)SECRETARY CONCERNED.—
15	reform	NDAA_FY2020.pdf	373	H . R . 6395—374 SEC . 838 .COMPTROLLER GENE
16	reform	NDAA_FY2020.pdf	412	H . R . 6395—413 an analysis of the process us...
17	reform	NDAA_FY2020.pdf	413	H . R . 6395—414 of the Department of Defense ...
18	reform	NDAA_FY2020.pdf	414	H . R . 6395—415 General of the United States ...
19	reform	NDAA_FY2020.pdf	543	H . R . 6395—544 ' ' (1) A staffing plan to ...
20	reform	NDAA_FY2020.pdf	551	H . R . 6395—552 ' ' (vi) transformation of ...
21	reform	NDAA_FY2020.pdf	553	H . R . 6395—554 (D) an assessment of the ex...
22	reform	NDAA_FY2020.pdf	568	H . R . 6395—569 (6) affirms the commitment ...
23	reform	NDAA_FY2020.pdf	580	H . R . 6395—581 (C)1 member selected by agr...
24	reform	NDAA_FY2020.pdf	581	H . R . 6395—582 intelligence services and str...
25	reform	NDAA_FY2020.pdf	587	H . R . 6395—588 of title 10 , United States C...
26	reform	NDAA_FY2020.pdf	589	H . R . 6395—590 involved in the illicit trade...
27	reform	NDAA_FY2020.pdf	928	H . R . 6395—929 Subtitle B—Military Family Ho...
28	reform	NDAA_FY2020.pdf	1112	H . R . 6395—1113 SEC . 4501 .OTHER AUTHORIZA

Writing simple table to file¶

Output File is in attached CSV at NDAA_FY2020_QuickSearch.csv

In [78]:

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

Further Analysis¶

In [ ]:

In [ ]: