5th Aug 2018 1:42PM

Task 1 - Collect Data from Web of Science.

Citation information for institutions

"Particular Domain" (Astronomy).

Particular institution (Inst 1)

- Citations received by Inst. 1 from other institutions (year wise) in csv file format

I_1

1990	1991	...	1995
10	12		6

→ citations received by Inst 1 from I2 in 1995 = 6.

I_2

1990	1991	- -	1995
1	6	- -	10

⋮

I_n.

Inst. 2 → And all the institutions which have cited Inst. 2

- arrange the citations year wise.

8 results- 92 + 73 + 32 + 28 + 18 + 17 + 7

= 267

Search for Result > Create Citation Report > In Citing Articles click "Analyze" > It will open up something saying "Showing 305 records for Total Citing Articles: OG = (Massachusetts Inst. of Technology (MIT)) and WC = Astronomy & Astrophysics

Analysis: Publication years: (2018)" > There in left menu click on "Organizations - Enhanced" > There you will get the details of all institutes that cited for that ~~topic~~ institute in that topic either in Treemap format or bar chart format (both .png files) that can be downloaded.

MIT

2014

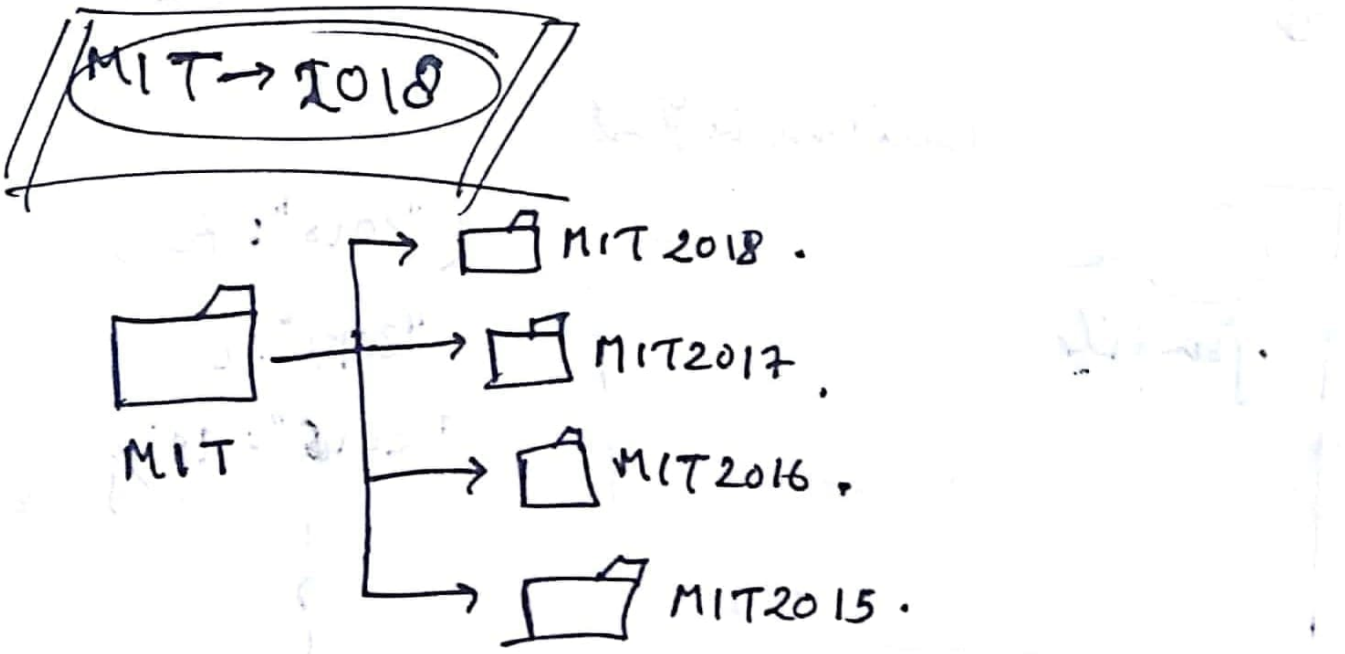
MIT : astronomy

J1

1994				2018
-				

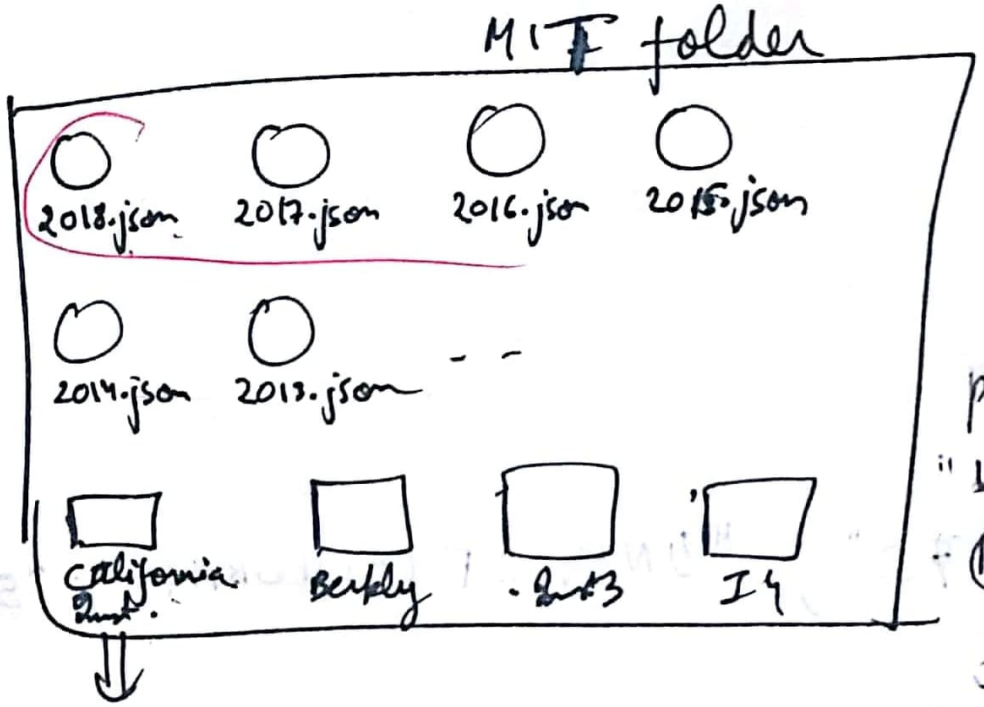
CSV

["MIT": "70", "UNIV OF CALIFORNIA": "56"]



MIT → MIT2018 → .json file having name & no. of citations received by MIT from various institutes in the yr. 2018.

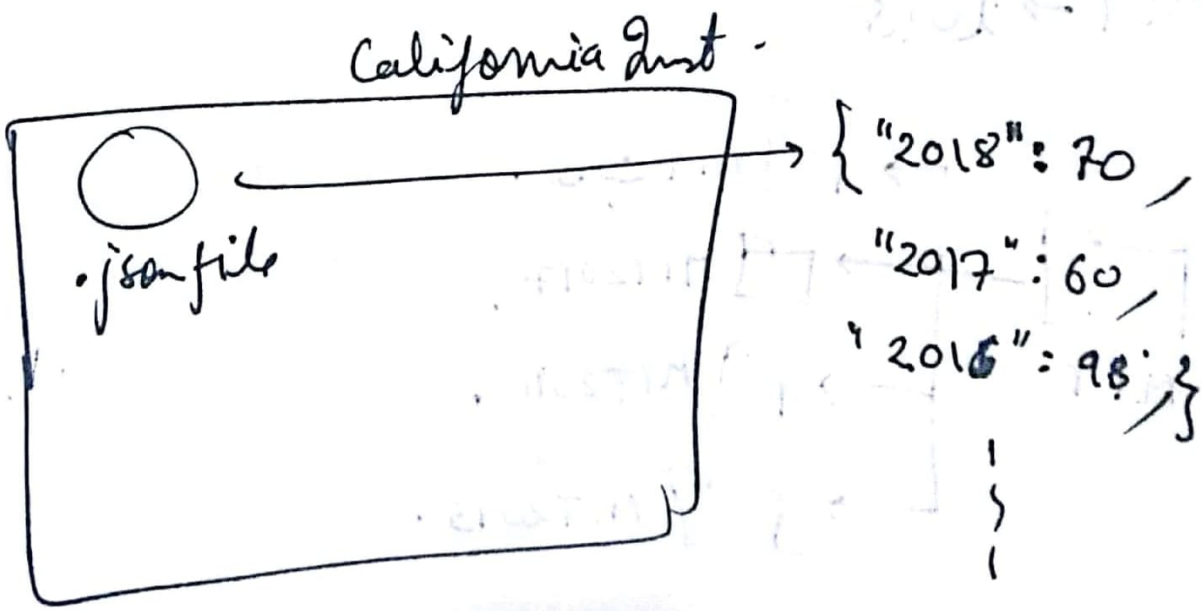
"LIT" : "70"



python3 scraper.py "url https://..."

MIT

folder name



1) Go to lib.lsu.edu

2) Login with mail & password.

3)

Goto the url given as →

Select the year you want the data for from the list. Ex→ 2018, 2017, 2016, ...

Now click Refine. Once done, click on "Create Citation Report" on the Right Hand Side menu.

Now click "analyze" under the "Citing articles" option. Now click on "Organization-Enhanced" option on left (on time option to select). Select "Record count" as 500 under the "Tree Map" visualisation (This is one time thing to do)

Now run the script 'scraper.py' with url (fixed url for all years) followed by folder name (preferably put it as name of Institute you are scraping data for) followed by year of which data you want.

Ex- python3 scraper.py "http://wcs.webofknowledge.com/RA/analyze.do?product=WOS&SID=6EdDEUNVu3VxOHR2NNP4&field=OG_OrgEnhancedName_OrgEnhancedName_en&yearSort=false" MIT "2017"

year.json (Ex- 2017.json) file will be made in MIT folder (Name of inst.) in the current directory. In this file just remove the last comma (a comma is extra due to which some error in json reading will come).

Once you get all the yearwise data in the files 2018.json, 2017.json, -- run the script test.py & it will create a folder "citing Institutes" inside MIT folder, inside which all the required citing institute files with year wise citation is present.