⊕ Coverage of open citations in DOAJ journals V.2

Constance Dami[1], Alessandro Bertozzi[1], Chiara Manca[1], Umut Kucuk[1]

[1]University of Bologna

2 ▾

Apr 21, 2022

dx.doi.org/10.17504/protocols.io.n92ldz598v5b/v2

Open Science 2021/2022

Chiara Manca

This is the protocol for the research of the coverage of open citations in DOAJ journals. Our goal is to find out:

- about the coverage of articles from open access journals in DOAJ journals as citing and cited articles,
- how many citations do DOAJ journals receive and do, and how many of these citations involve open access articles as both citing and cited entities,
- as well as the presence of trends over time of the availability of citations involving articles published in open access journals in DOAJ journals.

DOI

dx.doi.org/10.17504/protocols.io.n92ldz598v5b/v2

Constance Dami, Alessandro Bertozzi, Chiara Manca, Umut Kucuk 2022. Coverage of open citations in DOAJ journals. **protocols.io** https://dx.doi.org/10.17504/protocols.io.n92ldz598v5b/v2 Alessandro Bertozzi

citations, OpenCitations, DOAJ, open access, journals, open science

protocol ,

Apr 18, 2022

Apr 21, 2022

Apr 21, 2022 　Chiara Manca

60954

This protocol refers to a research done for the Open Science course of 21/22.

1   We download from the [DOAJ Public data dump](#) the article metadata in JSON format in order to collect all the article DOIs in DOAJ (doaj_articles_data.json). Then we open the file inside the python script, as indicated in the code below:

```python
import json

# Opening JSON file
f = open('doaj_articles_data.json')

# returns JSON object as
# a dictionary
doaj_articles_data = json.load(f)

# Closing file
f.close()
```

Once we have obtained the JSON file, we establish a connection with OpenCitations [REST API FOR COCI](#), using th*e* python library [request](#).

```python
import requests

r =
requests.get("https://w3id.org/oc/index/coci/api/v1/metadata/<DOI>"
)
```

Through the API we can collect all the field corresponding to the DOIs of the DOAJ articles. All the results of the request to the API are stored inside another JSON file (DOIs_API_OUTPUT.json).

```python
list_DOIs_API_OUTPUT = list()

for dois in doaj_articles_data:
    r =
requests.get(f"https://w3id.org/oc/index/coci/api/v1/metadata/{dois
}")
    list_DOIs_API_OUTPUT.append(r)

with open('list_DOIs_API_OUTPUT.json', 'w') as f:
    json.dump(list_DOIs_API_OUTPUT, f)
```

Then we divide the articles by the corrisponding journals.

1.1 We select and sum all the **citation_count** in the [metadata](#) of DOIs that come from each of the journals. We also collect for each journal all citations made to the articles of the journal.
The result will be a *citations*.**json** with all of the citations for each journal and **DOAJ_journals.csv** will be updated with a counter of the citations coming to the journals (**citations_count**).

1.2 We select and sum all of the [references_count](#) of DOIs that come from each of the journals. We collect for each journal all citations made by the articles of the journal.
The result will be a *references*.**json** with all of the citations for each journal and **DOAJ_journals.csv** will be updated with a counter of the citations done by the journals (**references_count**).

1.3 We check if the *citing* DOI in the citations of **references.json** and the *cited* DOI in the citations of **citations.json** is open by looking at the *oa_link.*
The result will be **open_citations.json** and **open_references.json** with all of the citations for each journal and **DOAJ_journals.csv** will be updated with a counter of the open citations coming and done by the journals (**open_citations_count**, **open_references_count**).

1.4 We combine all of the citations involving open access journals. We group them based on the *creation_date,* taking just the year, probably using a dataframe of the [pandas library](#) in python
The result will be **open_access_citations_by_year**.**json** of all of these citations for every year .that was found.

Data Visualization

2 We visualize our results with python libraries.
We specifically visualize **open_access_citations_by_year**.**json** in a graph.

Publishing data

3 We publish all the CSV and JSON data that we gathered.