Automated Collection of Scientific Publications Linked to NASA Earth Science Datasets



https://disc.gsfc.nasa.gov/

NASA/Goddard Earth Sciences Data and Information Services Center (GES DISC)

Irina Gerasimov^{1,2}, Itai Firan¹, Jerome Alfred^{1,2}, Armin Mehrabian^{1,2}, Mohammad Khayat^{1,2}, Binita KC^{1,2}, Long Pham¹, Jennifer Wei¹ and David Meyer¹

¹Code 619, NASA Goddard Space Flight Center, Greenbelt, MD, USA ²ADNET Systems Inc., Lanham, MD, USA

Abstract

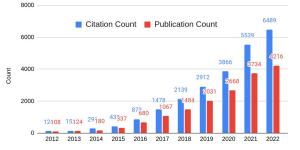
- The Earth Observing System Data and Information System (EOSDIS) began registering dataset Digital Object Identifiers (DOIs) in 2012.
- •As of January 2023, there are more than 11,000 dataset DOIs registered.
- By citing datasets by DOI, automated methods can be used to collect published works from a wide range of bibliometric databases.
- Presented is an automated workflow for collecting citations and methods for accessing the resulting library of citations.
- •Also presented are temporal trends of publication counts obtained from a variety of bibliometric sources.

Dataset DOI Citation Statistics [2012-2022]

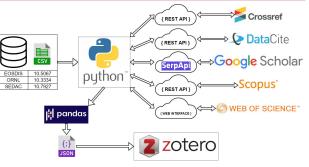
Methodology: Search Web of Science, Scopus, Crossref, Google Scholar and DataCite for publications referencing each of EOSDIS dataset DOIs. Retain only citations of books, journal articles, conference papers, dissertations, and reports (exclude Web pages, pre-prints, discussions, works without DOI, etc).

- Count of searched EOSDIS DOIs: 11,089
- Total count of unique publications found: 17,093
- Count of EOSDIS DOIs cited at least once: 3,012
- Total count of dataset citations in publications: 24,925

Total Dataset Citation and Publication Count by Year



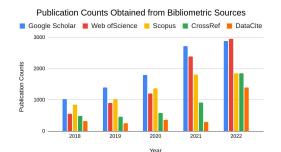
Publication Citation Collection Flowchart



In the last five years, the numbers of publications obtained from each bibliometric source show the following trends:

- CrossRef publication counts reached Scopus counts in 2022 increasing twofold compared to 2021.
- As compared to 2021, DataCite publication counts increased fivefold in 2022.
- For earlier years, Web of Science publication counts were lower than Scopus counts, but in 2021 they exceeded Scopus counts and in 2022 reached Google Scholar counts.

Bibliometric sources return a portion of unique publications not found in other sources. The total number of publications obtained in 2022 is 4,216, contributed by Google Scholar (2,832), Web of Science (2,887), Scopus (1,818), CrossRef (1,835) and DataCite (1416).

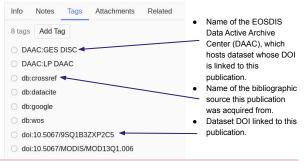


Accessing Citation Library from Zotero

Citation library is publicly accessible from an open source Zotero citation manager: https://www.zotero.org/groups/4567966/eosdis_dci/library

zotero	Groups Documentation Forur	ms Get Invo	lved
Other Group Libraries	⊥ [™] ≡ 101 ····		
REOSDIS_DCI	Title	Creator	Year
All DAACs	Catastrophic ice-debris flow in the Rishiga	Pandey et al.	2022
ASDC	Multi-spatiotemporal heterogeneous legac	Lian et al.	2022
C ASF	Updated trends of water management pra	Ranghetti a	2022
🗎 CDDIS	Development of earth observational diagn	Park et al.	2022
C GES DISC	Evaluation of MODIS-Landsat and AVHR	Rustanto an	2022
🗎 GHRC	Spatiotemporal inhomogeneity of total col	Park et al.	2022
C LAADS	Evaluating spatial accessibility to healthca	Jiao et al.	2022
C LP DAAC	A stacking ensemble algorithm for improvi	Zhang et al.	2022
C NSIDC	Estimation of the total dry aboveground bi	Migolet et al.	2022
C OB.DAAC	Risk assessment of rainstorm disasters in	Wang et al.	2022
C ORNL	Digital mapping of soil organic carbon stoc	Duarte et al.	2022
PO.DAAC	A method for reconstructing NDVI time-ser	Liu et al.	2022
C SEDAC	Mathematical vector framework for gravity	Hu et al.	2022

Zotero library content can be filtered by tags associated with each publication:



Future Work

Future work may include collaboration among EOSDIS DAACs on validation and expanding the citation library, depositing obtained linkages to DataCite, using the citation library for Al/machine learning and dataset discovery applications.