Harvesting Publications for Air Quality and Health Data

We are the Goddard Earth Sciences Data and Information Services Center (GES DISC), one of the 12 NASA Distributed Active Archive Centers (DAACs) that manage, archive, and distribute data as part of NASA’s Earth Observation Systems Data and Information System (EOSDIS). We provide support for the archive and distribution of the data from a range of satellite observations, models, ground measurements, and field campaigns.

- Multi-disciplinary data holdings include observations and model data of:
  - Atmospheric composition
  - Water/Energy cycles
  - Climate variability

Data are free and open to the public.

The Giovanni tool has a great impact in the scientific community as it has enabled more than 3000 peer-reviewed publications to date. The breakdown of variables used in Giovanni by publication is shown in the figure at right. Aerosol is the leading discipline used in these papers.

Air Quality and Public Health topics comprise 9% and 4% of 18,247 EOSDIS publications. For GES DISC, publications are gathered using Data DOI and short name.

GES DISC total count of publications is 11,064. Of those, 1,394 are Air Quality and 272 are Public Health. Shown below are counts of the citing of GES DISC datasets in these publications.

GES DISC is in the process of publishing Knowledge Graphs of publications using EOSDIS data in Kaggle Platform. The Knowledge Graph will help users to identify publications that are associated with a particular dataset that align with a science keyword.

Example: use of GES-DISC Knowledge Graph related to “Air Quality” keyword

- Blue nodes represent datasets
- Purple nodes represent publications that cite an example MERRA 2 dataset
- Red nodes represent Science Keywords