



The Benefits of Disseminating Data through the NASA Socioeconomic Data and Applications Center (SEDAC)

[SEDAC](#), the NASA Socioeconomic Data and Applications Center, is one of the Distributed Active Archive Centers (DAACs) in NASA's Earth Observing System Data and Information System (EOSDIS). Focusing on human interactions in the environment, SEDAC's mission is to develop and operate data and applications that support the integration of socioeconomic and earth science data and to serve as an "Information Gateway" between earth sciences and social sciences. SEDAC is managed by the Center for International Earth Science Information Network (CIESIN) at the Earth Institute of Columbia University. SEDAC is a member of the International Science Council (ISC) World Data System (WDS), a body that seeks to ensure best practices in data management and to contribute to a world where excellence in science is effectively translated into policy making and socio-economic development.

SEDAC disseminates many third-party data sets that are of use to the community of researchers studying human interactions in the environment. Examples include the Geocoded Disasters data set (Rosvold and Buhaug 2021), the Global Human Modification of Terrestrial Systems data set (Kennedy et al. 2019), and the Global Particulate Matter Grids (van Donkelaar et al. 2018).

Why Disseminate Data via SEDAC?

There are a number of reasons why disseminating data via SEDAC is beneficial to you, the author(s). First, disseminating your data set via SEDAC **increases its visibility** in the context of other relevant data, and is likely to **increase citations** of your data and associated journal articles or book chapters. SEDAC is visited by thousands of users each week and is widely recognized as a leader in human dimensions data, in particular spatial socioeconomic data. SEDAC encourages users to not only cite the data but to also cite the original work.¹ In addition, when appropriate, SEDAC generates open web services that are available to be used by SEDAC, NASA, and third-party clients.² Making your data available through these catalogs, services, and clients not only increases their visibility and ease of access, but also facilitates their integration with other complementary data in support of interdisciplinary research and applications.

A second reason to disseminate your data via SEDAC is to ensure their **long term preservation**. SEDAC retains multiple copies of all of its data holdings, with security copies in an offsite location. Although

¹ The journal article of reference is always found on the data set landing and download pages, and is included in accompanying documentation. SEDAC contributes metadata to the NASA Earth Data Search (<https://search.earthdata.nasa.gov/>), the Global Change Master Directory (GCMD) (<http://gcmd.nasa.gov/>), US government's data.gov catalog (<https://www.data.gov/>), the GEOS Data Portal (<http://www.geoportal.org/>), and DataCite services (<https://search.datacite.org/>).

² To view SEDAC's map client visit <http://sedac.ciesin.columbia.edu/mapping/viewer/>, and to view SEDAC's map services visit <http://sedac.ciesin.columbia.edu/maps/services>.



individual investigators often make good faith efforts to disseminate data via their institutional web sites, it is not uncommon for those resources to disappear after a period of time owing to staff turnover or web site updates. SEDAC assigns a persistent identifier (DOI) to each data set in its holdings. If the URL associated with a data set changes over time, the DOI will always point to the location of the data, even if data are eventually transferred to SEDAC's long term archive (LTA), which is managed in coordination with the Columbia University Libraries.

Another reason to disseminate your data through SEDAC is to receive **scientific peer-review and quality control** that evaluates and improves the quality of the data for reuse by a diverse audience of users. SEDAC reviews each data submission and describes each data product to enhance its potential to be discovered and reused by individuals interested in human-environment interactions. SEDAC may undertake, as well, conversion to additional data formats or changes to packaging of the data that will enhance its usability by the diverse group of users who are interested in human-environment interactions. Iterative data reviews are conducted by the SEDAC science team, by representatives of the SEDAC User Working Group (UWG) and NASA, and by the SEDAC Configuration Management Board (CMB). Upon approval for dissemination, data products are packaged with documentation, metadata and other relevant information or services to facilitate their utility to the scientific community, applied users, educators and students, and, where appropriate, the public.

Disseminating data via SEDAC also **contributes to the global scientific enterprise** through open data. There is abundant evidence that open data contributes to more rapid and equitable scientific progress and economic development. Furthermore, increasingly government funding agencies and major foundations require principal investigators to make their data openly available. SEDAC can help you to meet this type of requirement.

How to Disseminate Data via SEDAC

If you are interested in disseminating data via SEDAC, please take the following steps.

1. Visit the **Data Submission** page at <https://sedac.ciesin.columbia.edu/data-submission>. This page includes important information on the SEDAC's data curation criteria (see [Data Acquisition Criteria](#)). If you believe that your data set meets these criteria, please fill out the short [Data Set Submission Form](#). In that form, you will be asked to briefly describe the data set (subject matter, spatial extent, data formats) and to state whether it has been published in (or is in press or will be submitted to) a journal, or if there is some other form of documentation. If the data are already publicly available, you are asked to provide the URL.
2. SEDAC will respond to you within 1-2 weeks. If the data are of sufficient interest, we will request a review copy of the data to evaluate.

Should the data set remain of interest after internal review, a SEDAC staff member will work with you to put together a **data nomination package** for the SEDAC User Working Group (UWG), a group of 15-16 representatives from the scientific community who advise SEDAC on data and services that would be useful to the community. The UWG must approve the data set before it can be disseminated via SEDAC.



If the data are approved by the UWG, the next step is the **development of appropriate documentation**. If the data are already documented in a peer-reviewed journal article or book chapter that is open access, then it may be sufficient to package the data with the article (and any supplementary materials). If the journal article is not open access, then we would work with the authors to fill out SEDAC's standard documentation template. This documentation provides enough information for the users to be able to understand how the data set was constructed and how it may be used, as well as any limitations.

There are a number of internal procedures, from metadata creation to development of data set landing pages, archiving, peer-review, and quality control that generally require several months to complete. If you wish to reference the publication of the data via SEDAC in a journal article that is under review or in press, we are able to provide a data set landing page with a URL and digital object identifier (DOI) for the data set in advance of its publication.

We do hope that you will consider disseminating your data via SEDAC. Please visit the SEDAC web site at <http://sedac.ciesin.columbia.edu/> or contact SEDAC User Services at ciesin.info@ciesin.columbia.edu for further information about SEDAC and its services.

References

Kennedy, C. M., J. R. Oakleaf, D. M. Theobald, S. Baruch-Mordo and J. Kiesecker. 2019. Managing the Middle: A Shift in Conservation Priorities Based on the Global Human Modification Gradient. *Global Change Biology* 25(3): 811- 826. <https://doi.org/10.1111/gcb.14549>.

Rosvold, E.L., and H. Buhaug. 2021. GDIS, a Global Dataset of Geocoded Disaster Locations. *Scientific Data* 8:61. <https://doi.org/10.1038/s41597-021-00846-6>.

van Donkelaar, A., R. V. Martin, M. Brauer, N. C. Hsu, R. A. Kahn, R. C. Levy, A. Lyapustin, A. M. Sayer, and D. M. Winker. 2016. Global Estimates of Fine Particulate Matter Using a Combined Geophysical-Statistical Method with Information from Satellites. *Environmental Science & Technology* 50 (7): 3762-3772. <https://doi.org/10.1021/acs.est.5b05833>.