

## **JID DATASETS & BENCHMARKS ARTICLES Overview and Author Guidelines**

A scientific data descriptor article provides detailed information about a dataset, ensuring that the data is understandable and reusable by others. It should provide comprehensive insights for interpreting, utilizing, and reproducing data. Datasets may include images, text, video, and other modalities. Special consideration will be given to articles that describe multi-modal datasets, new datasets for large language models, clinical images, images obtained by in-vivo skin imaging techniques and dermatopathology scans, and images that include diverse skin tones and populations.

Benchmarks and Datasets Letters and Reports must connect to reputable repositories where the data, associated code, and processes are housed. Data should be valuable and relevant to dermatological research. The dataset's size should be sufficient, and it must have been approved, or granted a waiver of approval, by the author's Institutional Review Board or Research Ethics Committee, and written informed consent must have been obtained from all patients and control participants. A supplementary datasheet, as described by Gebru et al in the article "Datasheets for Datasets" is required (<https://arxiv.org/pdf/1803.09010.pdf>).

Datasets must be publicly available through open licenses that support modifications and the creation of derivative works. Repositories should ensure data permanency and quality curation; this includes assigning a permanent identifier to the dataset. Acceptable repositories include ISIC image archive (<https://www.isic-archive.com/>), Harvard dataverse [<https://dataverse.harvard.edu/>], and Hugging Face [<https://huggingface.co/docs/datasets/index>].

The reported dataset must be available to the editor and reviewers at the time of submission. The link must be used for the purposes of review only; it may not be used by the reviewer or editor for their own analysis or experimentation. Datasets must be made publicly available after acceptance.

- Letters should include the information listed below, without headings. Letters are limited to 1,000 words, 10 references, and 2 figures and/or tables.
- Original Articles should include the information below, with headings, and they should also describe the algorithms and models that were tested against the dataset.
- All Submissions should include a supplemental datasheet; for details see the article by Gebru et al "Datasheets for Datasets" (<https://arxiv.org/pdf/1803.09010.pdf>).

Include the following information:

1. Title
2. Abstract
3. Introduction  
Background and context, purpose and relevance of the dataset.
4. Methods  
Ethical review board approval or waiver thereof, data collection, statement of participants' written informed consent, instruments and equipment used, sampling methods, experimental setup and conditions), data processing (cleaning and preprocessing steps, data transformation and normalization; statistical methods, software and tools used.
5. Dataset Description  
General characteristics (size, types of data, etc.), variables and features, data quality.
6. Usage Notes  
Potential applications, limitations, best practices for using the dataset.
7. Data Availability Statement  
Repository Information including data and code, licensing and restrictions.
8. Conflict of Interest Statement
9. Acknowledgments
10. Author Contributions Statement (CRediT-compliant)
11. References
12. Tables
13. Figure Legends
14. Supplementary Datasheet  
Provide additional relevant information or data. See Gebru et al "Datasheets for Datasets" for details (<https://arxiv.org/pdf/1803.09010.pdf>).

Submissions may be made to JID at <https://mc.manuscriptcentral.com/jid>.

For more information, contact the Editorial Office at [JIDOffice@sidnet.org](mailto:JIDOffice@sidnet.org).