

# Datasheet

A datasheet is a standardized document used in order to provide information about the provenance, creation, and usage of a particular dataset. They serve to increase transparency and to make it easier for new users of the dataset to avoid unwanted biases. You can read more about Datasheets for Datasets at: <https://arxiv.org/abs/1803.09010>

Who created the dataset (e.g., which team, research group) and on behalf of which entity (e.g., company, institution, organization)?

No answer provided.

What do the instances that comprise the dataset represent (e.g., documents, photos, people, countries)? Are there multiple types of instances (e.g., movies, users, and ratings; people and interactions between them; nodes and edges)? Please provide a description.

No answer provided.

What data does each instance consist of? “Raw” data (e.g., unprocessed text or images) or features? In either case, please provide a description.

No answer provided.

Are there any errors, sources of noise, or redundancies in the dataset? If so, please provide a description.

No answer provided.

Is the dataset self-contained, or does it link to or otherwise rely on external resources (e.g., websites, tweets, other datasets)? If it links to or relies on external resources, a) are there guarantees that they will exist, and remain constant, over time; b) are there official archival versions of the complete dataset (i.e., including the external resources as they existed at the time the dataset was created); c) are there any restrictions (e.g., licenses, fees) associated with any of the external resources that might apply to a future user? Please provide descriptions of all external resources and any restrictions associated with them, as well as links or other access points, as appropriate.

No answer provided.

Does the dataset contain data that might be considered confidential (e.g., data that is protected by legal privilege or by doctor-patient confidentiality, data that includes the content of individuals’ non-public communications)? If so, please provide a description.

No answer provided.

What mechanisms or procedures were used to collect the data (e.g., hardware apparatus or sensor, manual human curation, software program, software API)? How were these mechanisms or procedures validated?

No answer provided.

Were any ethical review processes conducted (e.g., by an institutional review board)? If so, please provide a description of these review processes, including the outcomes, as well as a link or other access point to any supporting documentation.

No answer provided.

How can the owner/curator/manager of the dataset be contacted (e.g., email address)?

No answer provided.

Any other comments?

No answer provided.