AWS questionnaire:

* Contact name.
* Contact email address and link to GitHub issues page (if relevant).
* Name of institution.
* Website of institution.
* Name of dataset.
* Brief description of dataset (about 200 words to describe the data, how it is created, and what it is used for).
* Size of dataset (e.g. 100GB).
* Growth rate of dataset in GB per year, if applicable (e.g. 50GB per year).
* Dataset file formats (e.g. Parquet, ORC, csv, cloud-optimized GeoTiff, NetCDF, json).
* Dataset license (e.g. Creative Commons, no license, US Government work).
* Explain how will this data could be used (e.g. training machine learning models, climate research, computer vision, energy efficiency, genome analysis, training, agricultural research, demo development. Specific tools or workflows would be very helpful).
* What’s new or novel about this dataset or the way it is provided? (Tell us why this dataset is awesome).
* How dataset has been optimized to be used in the cloud.
* Other DOE labs, institutions or companies already using this data.
* URLs pointing to documentation of the structure and content of the dataset (attaching a readme file is acceptable if no such link exists).
* Do you have any tutorials or example data usage materials (i.e. Jupyter Notebook) that can be modified to access the data via AWS? If so, please include a URL or attach a file.
* How NREL intends to promote the existence of this dataset on AWS, including blogs, mailing lists, or other channels.
* Tags (chosen from this list: <https://github.com/awslabs/open-data-registry/blob/main/tags.yaml>. If you’d like to add tags that aren’t on this list, notate them as “not in list”).
* URLs to any papers or reports about this dataset.