

## Licenses

We adhere to [Apache 2.0 License](#) from Aya Dataset and Aya Collection and [Terms of Use](#) for GPT-4 when constructing our **M2Lingual** dataset. We confirm that we bear the responsibility in the case of violation of rights and will take appropriate course of actions if needed. Our dataset is licensed through CC-by-NC-SA-4.0 license. The dataset will be hosted on HuggingFace datasets and maintained by the authors.

## Dataset Documentation

Dataset documentation and sub-samples are available at <https://huggingface.co/datasets/ServiceNow-AI/M2Lingual>. The Croissant metadata associated with the dataset is available at <https://huggingface.co/api/datasets/ServiceNow-AI/M2Lingual/croissant>. It is intended to be used to improve model performance towards multilingual natural language understanding.