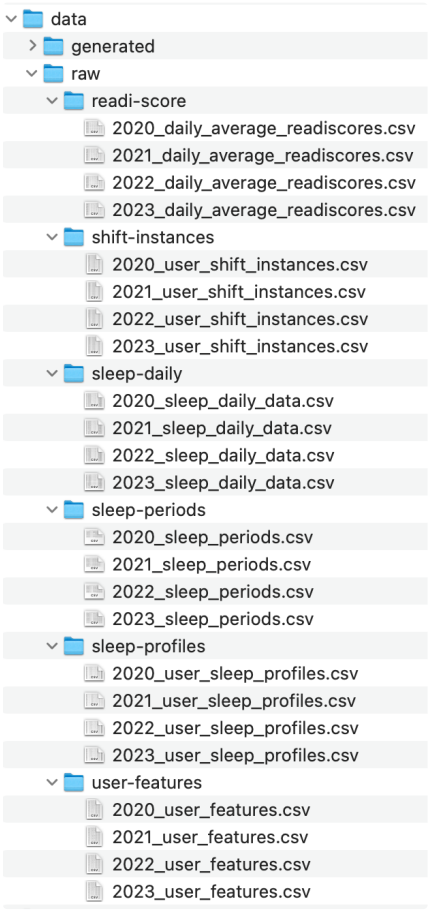
Steps to run the code (fatigue-prediction package version 1)

1. Ensure you have Python installed on your system.
2. In the `data/raw/` directory, add the data for each year (2020, 2021, 2022, 2023) in the respective subdirectories (with same name of directory and files as shown in the image).



Directory structure for raw data files

3. Create a new virtual environment and install the required dependancies. You can find the documentation to do the same here, How to create virtual environment using pip and install the required libraries?
4. The **input_user_list.txt** file will be available in the `config/` directory. Add the list of users, separated by commas, for whom you want to train the models, predict the results, or do both, to the file.
5. The **input_config.yaml** file will be available in the `config/` directory. Provide the start_date in *YYYY-MM-DD* format and total_days_to_predict. In order to make predictions, these parameters will be used.
6. Run model training file using `python main_training_pipeline.py` command. This will iteratively create sleep minutes, sleep start time models (for day and night) for each user in the **input_user_list.txt** file. This program will extract, preprocess and train the model for the user provided. It will chronologically execute main_extraction_pipeline.py, main_preprocessing_pipeline.py and main_modelling_pipeline.py.
   Run `python main_extraction_pipeline.py` to only extract data from the dataset dump.
   Run `python main_preprocessing_pipeline.py` preprocesses user data and generates features for modelling.
   (NOTE: If `python main_training_pipeline.py` is executed, there is no need to run `python main_extraction_pipeline.py` and `python main_preprocessing_pipeline.py` commands)
7. Run the prediction file to predict the sleep timings using `python main_prediction_pipeline.py` command. This will iteratively predict the results for all the user provided in **input_user_list.txt** based on parameters provided in **input_config.yaml** file.

8. The results will be produced in result.csv file. It will have user_id, sleep_start_time, sleep_end_time for all the users provided in **input_user_list.txt** based on parameters provided in **input_config.yaml** file.
