

# RawImgMetaData Usage Instructions

This document outlines how to configure and run the metadata generation script using the provided input files and configuration settings.

## 1. Prepare the Configuration File

Ensure that a `config.json` file exists in the root directory of the project. This file defines all the necessary static metadata and file paths used by the script. Below is an example structure:

```
{
  "city": "Watkinsville",
  "state": "GA",
  "Standard_Title": "The Oconee Enterprise",
  "LCCN": "sn89053667",
  "Edition_Order": "1",
  "Page_Physical_Description": "Print",
  "Reproduction_Agency": "Digital Library of Georgia; Athens GA",
  "Reproduction_Agency_Code": "gu",
  "Reproduction_Note": "Present",
  "Physical_Location": "Georgia Newspaper Project; Athens GA",
  "Physical_Location_Code": "gnp",
  "Batch_Name": "oew2023",
  "Reel_Number": "oew2023",
  "input_file_name": "Raw_OEW_2023.xlsx",
  "output_file_name": "OEW_2023_metadata.xlsx",
  "input_path": "/path/to/input/folder",
  "output_path": "/path/to/output/folder"
}
```

## 2. Prepare the Raw Input Excel File

The input Excel file must be placed in the folder specified by the `input_path` key in the `config.json`. It should include the following required columns:

- `Date` (in datetime format)
- `Total_Pages` (integer)
- `vol` (volume number)

- `issue` (issue number)

This file must match the expected structure exactly—**no additional columns, no missing columns, and no mismatches in column names or capitalization**. Any deviation will cause the script to fail or behave unexpectedly. are correctly capitalized and formatted to match what the script expects.

### 3. Run the Script

Once the project is properly configured and the input file is ready, run the script using PyCharm or from a terminal:

```
python main.py
```

### 4. Output

After execution, the script will generate a new Excel file in the location specified by the `output_path` in your configuration. This file will be named according to the `output_file_name` value (e.g., `OEW_2023_metadata.xlsx`) and will include fully formatted metadata across 21 columns, one row per page.

You should see a confirmation message like:

Excel file saved successfully to: `/your/output/path/OEW_2023_metadata.xlsx`

Make sure to review the output file for correctness and completeness based on your project requirements.