**Step-by-Step-Guide**

**How to Use This MUST GitHub Project**

**1. Requirements**

* **Installed Software:**
  + [R (Version 2022.02.3 or higher)](https://cran.r-project.org/)
  + [RStudio (optional but recommended)](https://posit.co/download/rstudio-desktop/)
  + Required R packages: dplyr, openxlsx, purrr, stringr.

**2. License**

This project is licensed under the MIT License – see LICENSE.txt for details.

**3. Clone or Download the Project**

* **Option 1: Clone via Git (Recommended)**  
  Open a terminal or Git Bash and run:
* git clone <https://gitlab.lrz.de/cccm/must_private/>.git

Then navigate into the project folder:

cd [must\_private]

* **Option 2: Download as ZIP**
  + Go to <https://gitlab.lrz.de/cccm/must_private/>
  + Click **„Code“ → „Download ZIP“** and extract the files.

**4. Project Structure Overview**

[must\_private]/

├── MUST.R # Main script for initiating the analysis

├── input\_data/ # Directory for input data (file data\_tnm\_full.csv must be provided by the user)

├── mapping\_tables/ # Mapping files required by the main script

├── copyright\_catalogs/ # Directory for user-defined mapping files (files must be provided by the user, following the example files)

├── scripts/ # Additional R scripts required by the main script

├── results/ # Directory for analysis outputs (will be created by the main script)

├── step\_by\_step\_guide.docx # User guide with detailed execution instructions

└── README.md # General project information and usage notes

**5. Prepare Input Data**

* Place your input data files in the input\_data/ folder
* Ensure filenames and formats match the expectations defined in the paper.

**6. Run the Project**

* Open R or RStudio.
* Set the working directory to the project folder: setwd("must\_private")
* Run the main script: source("MUST.R")

**7. View Results**

* All outputs will be saved in the results/ folder.
* Generated tables will be in standard formats like .csv, .xlsx

**8. Troubleshooting**

* Ensure all required R packages are installed. You can install them by running:
* install.packages(c("dplyr", "openxlsx", "purrr", "stringr"))
* Check if the input files are correctly formatted and placed in the right folder.