

Preparations for SICSS session on "Using LLMs to Create Data Analysis Pipelines for Text-as-Data Research"

Seraphine F. Maerz (University of Melbourne)

Welcome! Please complete these setup steps **before the workshop** so you can jump right into the exercises. **Estimated time:** ~15 minutes (plus download time). **New to R?** No worries — everything is explained for beginners.

Step 1: Install R and RStudio.

R is the programming language we'll use. RStudio makes working with R easy. **Install R:** Go to <https://cran.r-project.org/> and click the download link for your system (Windows or Mac). Run the installer with default settings. **Install RStudio:** Go to <https://posit.co/download/rstudio-desktop/>, download the free version, and install it. Open RStudio to verify it works.

Step 2: Create a Project.

In RStudio, click **File** → **New Project** → **New Directory** → **New Project**. Name it **quallmer** (or whatever), choose a location, and click **Create Project**. Always open this **.Rproj** file when working on workshop materials.

Step 3: Install R Packages.

In RStudio, look at the **bottom-right pane**, click the **Packages** tab, then click **Install**. Type: **tidyverse**, **quallmer**, **rollama**, **usethis** and click **Install**. Wait for installation to finish, then verify by typing this in the **Console** (bottom-left pane) and hitting **Enter**:

```
library(tidyverse)
library(quallmer)
library(rollama)
```

Step 4: Install Ollama and Download a Model.

Ollama lets you run open-source AI models locally and for free — no fee-based API keys needed. **Install Ollama:** Go to <https://ollama.com/>, click **Download**, install, and open the app (runs in the background). **Download the model:** In the RStudio **Console** (bottom-left pane), type this and hit **Enter**:

```
library(rollama)
pull_model("llama3.2:1b")
```

This downloads a ~1.3 GB *lightweight* model. Once complete, test with `ping_ollama()` — you should see a confirmation message.

Step 5: Set Up OpenAI API Key.

I recommend setting up an OpenAI API key to experiment with commercial models. We will discuss risks and precautions when working with commercial models during the workshop. They are not for free, but you can set a spending limit to avoid unexpected charges. For the workshop, we will only need a few cents worth of API calls, so the \$5 limit is more than enough. You can also use the free Ollama models if you prefer. **Important:** The OpenAI API is *different* from a ChatGPT account — you need a separate API account.

1. Go to <https://platform.openai.com/> (not chatgpt.com) and sign up
2. Add a payment method under **Settings** → **Billing** and set a \$5 spending limit
3. Go to **API Keys** → **Create new secret key** → Copy the key

Store it in R by running `usethis::edit_r_environ()` in the Console, then add this line to the file:

```
OPENAI_API_KEY="sk-your-key-here"
```

Save, close, and **restart RStudio**. Always keep your API key private and never share it. If you accidentally expose it, revoke the key immediately from the OpenAI dashboard.

You're ready :-) See you at SICSS. Questions or issues? Please email me to: seraphine.maerz@unimelb.edu.au