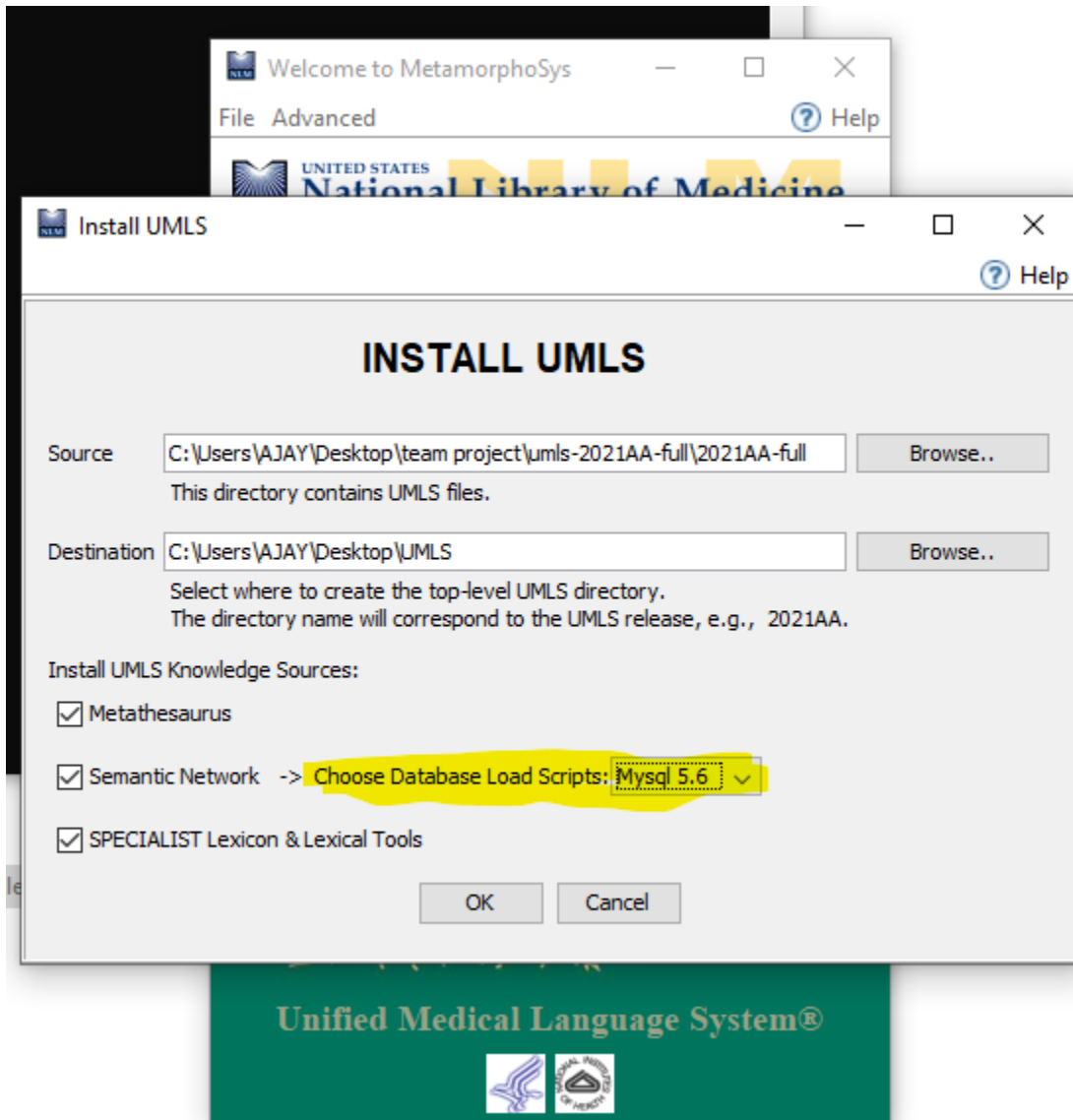


1. Download UMLS full from
<https://www.nlm.nih.gov/research/umls/licensedcontent/umlsknowledgesources.html>
2. Run the **run64** batch file(**run_mac** for mac) and select Choose Database Load Scripts
Mysql 5.6



3. Install MySQL 5.5 and follow the instructions at
https://www.nlm.nih.gov/research/umls/implementation_resources/scripts/REALDME_Semantic_Network_MySQL.html
Create UMLS database using below command-
CREATE DATABASE IF NOT EXISTS umls CHARACTER SET utf8 COLLATE utf8_unicode_ci;
4. Go to the UMLS installation directory(in the screenshot above its
C:\Users\AJAY\Desktop\UMLS) and navigate to \2021AA\META

Inside this folder you should see the **populate_mysql_db** script. Edit this script and set the below variables with your sql path and credentials-

set MYSQL_HOME="C:\Program Files\MySQL\MySQL Server 5.5"

set user=root

set password=password

set db_name=umls

Once these credentials are updated run the batch/shell script.

5. Edit the **KnowledgeGraphGenerator.py** and edit the sql credentials from line no 18

```
#database config values. Please edit these variables before executing the annotated file
host_name = "localhost"
database_name = "umls"
user_name = "root"
password = "password"
```

Note: This script assumes that the annotated csv file is in the same directory as the **KnowledgeGraphGenerator.py** script. If not, the paths need to be changed in the script.

This script will generate fifty files - eg.- annotations_pubmed21n0001_parsed_triples.spo

This is all of the files that need to be added to the brain/kgs/ directory of the kbert_tp solution in the github

folder. (https://github.com/marlenakuc95/BASF-team-project/tree/main/kbert_tp)

All of the SPOs are also available on

https://drive.google.com/drive/folders/1og1PT14opUv6W89_kqBETn CZq9KF_P2t?usp=sharing