**MetaMap wrapper project**

*Written by Jonathan Shifman  
January 2018*

**General Description**

MetaMap is a tool for mapping various medical names and concepts to a smaller set of standard concepts, to allow categorization and easier analysis of entire medical texts.

This project aims to provide a wrapper (written in python) for the MetaMap functionality. The project makes use of an existing MetaMap wrapper module, called **pymetamap**, while manipulating its output to produce results in a certain (and configurable) format.

The project source can be found in this [GitHub repository](https://github.com/JonathanShifman/metamap-project).

The code of pymetamap has been included in the repository as a submodule.

**Set-up**

To use the wrapper, you will need to perform the following steps:

Download & Install MetaMap

Download the latest version of MetaMap from the [downloads page](https://metamap.nlm.nih.gov/MainDownload.shtml).  
Make sure to download the Linux release, since the project is currently not compatible with windows.

Follow the installation instructions to install MetaMap. Be sure to start the Tagger server as described.

Configure the wrapper to use the MetaMap binary

In the config.ini file, set the meta\_map\_path variable to point to the binary of Meta Map you installed.

Set the relevant\_field\_names variable to contain the fields you want included in each concept dictionary.

Some possible field names:

score – MetaMap indexing score  
preferred\_name – The preferred name of the concept  
cui – The concept’s unique identifier  
semtypes – Semantic types list

After having performed these steps, you can start using the wrapper.

**Usage Instructions**

Initialize an instance of MetaMapWrapper, like so:

*wrapper = MetaMapWrapper()*

To analyze a specific sentence, invoke the following method:

*output\_dict = wrapper.analyze\_sentence(sentence)*

Alternatively, you can analyze multiple sentences at once by invoking:

*output\_dict = wrapper.analyze\_sentences(sentences)*

In both cases, output\_dict is a dictionary mapping serial indices to dictionaries, each representing a concept tagged by Meta Map.

A concept dictionary contains a mapping from its field names to their values.

An example usage of the wrapper, demonstrating the output data structures, can be found in ExampleUsage.py.