LBD\_SemMed.py is a python script which visualizes the relation between gene, drug and disease. It takes two concept names as input (either drug name, gene name or gene name, disease name or drug name, disease name), and returns the concepts of the third type which the two input concepts are associated with and represent them in the form of network graph, where vertices are concepts and edges are predicates. Blue is disease, yellow is gene and green is drug.

1 Users should install python 3.5.2 on their laptops.

2 Please use pip3 to install networkx, matplotlib, numpy and elasticsearch-py with the following commands:

$ pip3 install networkx

$ pip3 install matplotlib

$ pip3 install elasticsearch 2.x.y

$ pip3 install numpy

4 User should download Semantic Medline databases PREDICATION\_AGGREGATE table from <https://skr3.nlm.nih.gov/SemMedDB/index.html> , load to local MySQL server

5 Users should install elasticsearch 2.x, and logstash (elasticsearch Xpackage)

6 Users should install logstash jdbc by following command in logstash bin folder:

bin/plugin install logstash-input-jdbc

( to use jdbc, download mysql-connector-java-5.1.39-bin.jar from <https://dev.mysql.com/downloads/connector/j/> and put it into logstash folder)

Then use logstash to log PREDICATION\_AGGREGATE table into elasticsearch index. Create an index named “semantic\_medline” and doc type “predication\_aggregate”.

LS\_HEAP\_SIZE=4g ./logstash –f {config file name}.conf

(About how to write a configure file please read sample\_config.conf. )

7 before run LBD\_SemMed.py, please run elasticsearch

./elasticsearch