Python files:

1. Dataret.py (for retrieving documents from Wikipedia)
   1. Input
      1. “nwdb” (list of key words)
   2. Output
      1. “All” folder with Wikipedia contents
2. Cooc\_countnet.py (for generating file keyword map)
   1. Input
      1. “nwdb”(list of keywords)
      2. Documents in “all” folder with filenames as in “nwdb”
   2. Output
      1. Filewords(om)3 – this has which keywords present in which file
3. Matcre.py (co-occurrence graph)
   1. Input
      1. “Filewords(om)3”
      2. “nwdb”
   2. output
      1. “cooc.gdf” co-occurrence graph
4. Wddistime.py (word-distance graph)
   1. Input
      1. “filewords(om)3”
      2. “nwdb”
   2. Output
      1. Wd\_timeline.gdf- word distance graph
5. Sptree.py (spanning tree generation from gdf)
   1. Input
      1. “sptreedata” – list of edges present in graph (in GDF)
      2. nwdb
   2. Output
      1. “spfinal.gdf”
6. Combcowd1.py (combination of co occurrence and word distance)
   1. Input
      1. “wd” file with list of edges from word distance gdf file
      2. “cooc” file with list of edges from co occurrence gdf file
      3. “nwdb”
   2. Output
      1. Comb1.gdf