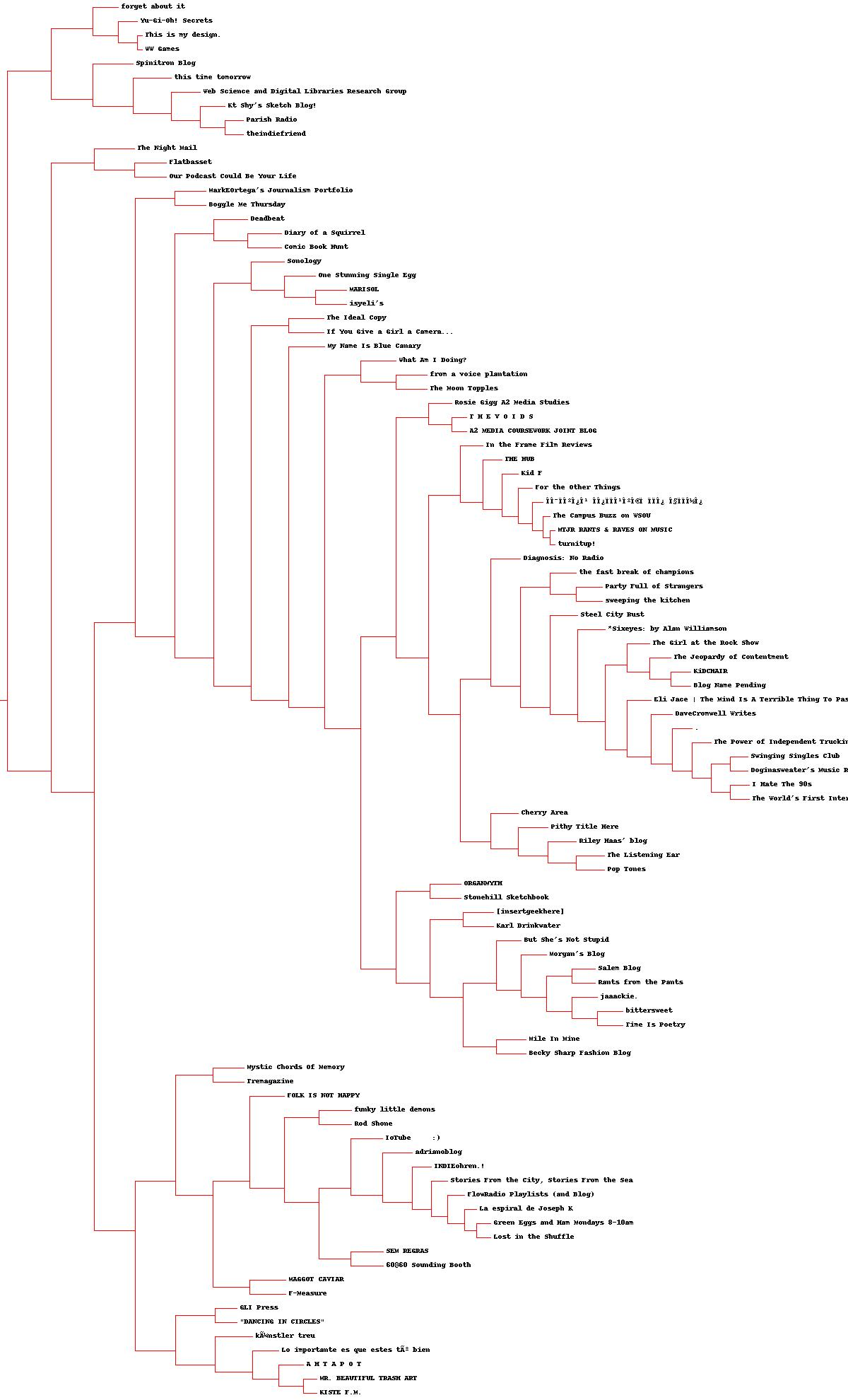
CS432 Assignment 8

1. Four Python scripts, fetchFeeds.py, generatefeedvector.py, clusters.py and eliminatedWords.py, were used from <https://github.com/shawnmjones/cs595-f13/tree/master/assignment9>. Clusters.py is necessary for generatefeedvector.py and eliminatedWords.py to work. Clusters.py and generatefeedvector.py originally came from Toby Segaran. Feedlist.txt was created from fetchfeeds.py and then feedlist.txt was inputted into generatefeedvector.py to generate the data for blogdata1.txt. Blogdata1.txt was then inputted into eliminatedWords.py with 500 as an argument and data.txt was created as a result.
2. A modified version of makeDendogram.py from Shawn Jones was used to create the dendogram image (‘blogclust.jpeg’) and the ascii file (‘ascii-dendogram.txt’). Clusters.py is required to allow this script to work.

[This part left intentionally blank.]



1. A modified version of makeClusters.py from Shawn Jones was used to generate the K-Mean list (‘KMean.txt’), the number of iterations each value of *k* produces. Clusters.py is required again to allow this script to work. K=5 has 7 iterations, K=10 has 8 iterations and K=20 has 4 iterations.
2. A modified version of makeMDS.py from Shawn Jones was used to create the MDS image (‘MDS.jpeg’) and the MDS 2D Coordinates file (‘MDS.txt’). It took 281 iterations to create the MDS.