CS 4395.001 - Assignment 6: Report

I created the knowledge base by simply going through compiling a dictionary with each of the top 10 most frequent words as the key. Then, each line of each text file is then iterated through to find every sentence containing the one of the top 10 words. A list of these sentences is created, and it becomes the value for each word (the key) in the dictionary.

Screenshots of the knowledge base:

The Top 10 most frequent terms in all 15 sites are:
1. anime
2. evangelion
3. rei
4. shinji
5. like
6. series
7. characters
8. one
9. many
10. genesis

A sample dialog that a chatbot using this knowledge base could use to hypothetically have:

User: What is the name of this anime?

Chatbot: The name of this anime is Neon Genesis Evangelion

User: Who is the main character?

Chatbot: The main character is Ikari Shinji.

User: Who are some of the supporting characters?

Chatbot: The supporting characters are Rei, Asuka, and Misato.