The chatbot I built is based upon the character Sheldon Cooper from the hit TV show The Big Bang theory is a comedy show that revolves around the lives of four scientists in their daily lives. Arguably the most iconic character was Sheldon Cooper. Sheldon Cooper quickly became a fan favorite in the show due to the uniqueness of his character on TV. The character is well scientific mindset, extremely type A personality, and his social ineptitude often leading to complications. The idea I had was to build a chat bot to emulate Sheldon Cooper.

one of the most difficult aspects of this project was the data collection. The first hurdle was finding out how I could get enough data about Sheldon Cooper to create a chat bot based upon him. So to start I need to find a way to get the script for every Big Bang Theory episode. The first place I looked was Cagle. Cagle is a site with the millions of datasets in every topic. When I checked on kaggle I found one data set with Big Bang Theory scripts, however upon further inspection the data set only had a scene or two from each episode Rather than the entire episode. So with that data set I wouldn't have had enough data. The next place I checked was Google. On Google I found a script sites. The script site had the scripts for every episode, however it was on the web page, not in PDF or any other easy manner formatic inputs. So to build up ideas of web the Wiz is age show 25 episodes of time period then once you click on one of the episodes it would take you to a page with the entire script for that episode. The first thing I wrote was code to grab the script for an episode when given a link. The way I did this was by looking for the paragraph tags in the web page and then getting the text from those paragraph tags. In each of the paragraph tags the site had one line or one stage direction from the episode. So I put all that text into a list.