**Project Prospectus**

Sam Yuxiao Sun

1. **Topic**

I would like to explore textual contents of Twitter and Sino Weibo (Chinese version of Twitter) and conduct a discourse analysis. In specific, I would like to see and compare how national image of US is perceived by Weibo users in China, and how that of China is perceived by Twitter users from US. In particular, it would be interesting to see how these images are related with government propagandas, diplomatic strategies and bilateral relationship. Additionally if I can retrieve data on users’ age cohorts, I will identify and compare the generation difference within and between the two countries. I may narrow down my focus to more specific issue depending on my progress.

1. **Data and Code**

For Twitter data I will use the tweepy module. One challenge is to design proper filters to censor tweets from us users only. As for Weibo, there are similar modules available and I have found a good scrapping tutorial from the webpage <http://chengjun.github.io/en/2014/03/scraping-weibo-using-python/>. An complication might arise due to the nature of Chinese language in terms of words parsing <http://www.cs.cmu.edu/~wcohen/postscript/emnlp-2014.pdf> and I am reading more literatures to decide my best strategy.

1. **Team Member**

Single person group.