Project Proposal

**Group Members:**

Gopal Menon and Kimberly Williamson

**Data Sources:**

News articles from various news outlets. We are looking at the following APIs to gather the news articles:

* Event Registry (This includes the body of the article, but the amount of articles available might be less than the News API)
* News API (This only gives a description, so we would need to find a way to crawl the websites when given an URL)
* AlchemyData News (Requires IBM Bluemix account. There is a free 30-day trial, but we might be limited to collecting all of the articles in 30 days.)
* RSS Feeds (This only gives a description, so we would need to find a way to crawl the websites when given a URL)

**Structure:**

We are proposing to group similar news articles. We will then look at the groups independently to determine the various spins of the similar grouped articles. (Add information about clustering?)

**Problem:**

Given the recent visibility of Fake News, we believe it would be interesting to understand the patterns in how news is reported. By looking at various news outlets, we hope to label the biases that are added to articles containing similar factual content. While this will not change the spread of Fake News, we hope to better understand patterns related to the subjectivity added in the reporting of stories across sources.

**What is New:**

I don’t know what to add here.