**News Aggregator**

**Problem**: In today’s world, there is a lot of concern over whether news outlets are biased or not, and whether there is selective reporting. However, it may not occur to all consumers that their news sources are biased, so there is a problem of not knowing what has never been shown. For instance, there was a lot of criticism in the recent election towards Facebook for how they selectively reported news to its users. This is something that should not run rampant in today’s society.

**Solution/Project**: There should exist a news site that aggregates the topics being talked about in every community and in every region with the purpose of pointing out those topics that are limited to certain regions or organizations, and those topics that present across many regions and many organizations. The goal is to break the aforementioned information blind spot, something that is especially dangerous in a technologically advanced society.

**How Rampant is the Problem**: For nearly every single person in the Western world, there exists an information blind spot. Limiting information intake from a few selected sources is the ultimate cause. Even communities such as Reddit face this problem, where the user base has a very common personality or interest base, and thus only information that the collective mind accepts will be allowed to be disseminated. There exists a major problem with many mindsets, even if a tool like this was built: Willful ignorance. One can only hope to create an environment in which unbiased information distribution exists.

**Steps to Project Completion**:

1. Create a diverse list of cities that represent a wide spread set of geographical locations. Gather information such as region, state, population, and average income as supplementary data.
2. Create a list of newspapers and media sources that exist within these cities
3. Create a general scraping code that will suck in the relevant pages from the media sources
4. Use NLP and Bayes to determine the topics of any given media source and the probability that other media sources have the same topic
5. After determining which media sources have similar topics, if there are enough sources for a given topic, provide the link to the most credible sources (this definition is still pending)
6. Create a set of traits by which the measure the media sources; the goal is to find the distribution of sources across each trait, and to rate each media source
7. Oliver, the rest is you