

### 1. Topic modeling and text summarization and for new articles

**Idea:** There is abundant amount of news being generated everyday, to help users to select the news articles they want to read into, a quick summarization of each new article can be very useful. Beside that, use NLP to group all the news into various topics and indicate relevance between different news due to characters or regions, so that user can select their own interested topic, or find relevant news article according to the ones they have already read.

**Data: All the news** - 143,000 articles from 15 American publications

<https://www.kaggle.com/snapcrack/all-the-news>

### 2. Topic modeling and text summarization for the gospels

**Idea:** There are four gospels, Matthew, Mark, Luke and John, in the Bible, each records the stories and teaching of Jesus. We can use NLP to find out the key individuals and central topics in each of the gospel, and discover the topics that have being covered in all the gospels, and also the ones that are unique to only one or two.

**Data: Bible Corpus**-English Bible Translations Dataset for Text Mining and NLP

[https://www.kaggle.com/oswinrh/bible#key\\_abbreviations\\_english.csv](https://www.kaggle.com/oswinrh/bible#key_abbreviations_english.csv)

### 3. Topic modeling and text summarization for the Epistles of Paul

**Idea:** Apostle Paul wrote several letters to early churches and disciples, Like Philippians, Ephesians, Corinthians, Timothy, for passing on the teaching of Jesus. We can use NLP to find out the topics in each of the letter, and discover the topics that have being covered in all the gospels, and also the ones that are unique to only one or two. After all, we may find the epistles that are most similar in terms of content.

**Data: Bible Corpus**-English Bible Translations Dataset for Text Mining and NLP

[https://www.kaggle.com/oswinrh/bible#key\\_abbreviations\\_english.csv](https://www.kaggle.com/oswinrh/bible#key_abbreviations_english.csv)