

CS 416 Data Visualization – Create A Narrative Visualization Project

Rakesh Patnaik

r_patnaik@illinois.edu

1. Introduction and Hosted Location

A visualization of news headlines over the years 2017 - 2021 that helps analyze key trending geopolitical activity and other drivers during this period. This visualization uses a dataset from ABC news and its top headlines. This dataset was further curated to limit to years 2017 - 2021. Natural language processing was used to extract Person Names, Location Names, Organization Names and Nouns or words describing things. Visualization is hosted at the following location

- <https://rakesh-patnaik.github.io/cs416narrativevisualisation/>

2. Messaging

News headlines are key to events and happenings in the past. They leave a trail of information at a high level that enables analysis and trending during a retrospect in years to follow. The idea of this visualization is to provide the users with a high level graphic view of words that defined news during the period of 2017 - 2021. With the capability to identify key words that defined headlines during a particular year user can then use the information and compare over years to draw trends. Trends such as the rise and cure of COVID-19 in 2020 and 2021 , Russian sanctions of 2017 and 2018 are some of the trends that one can derive.

3. Narrative Structure

Narration includes a interactive slide show structure. A default view of top keywords in the year 2021 is shown and the user is presented with options to navigate to other years or to open a view to compare top keywords across the years. Options to further filter by Person Names, Organization Names and Location Names helps drive curiosity on specific keywords that the user would be interested, and hence pick to further draw trends year-over-year using the comparison view. Different audience might end up using a option that best suits their curiosity and area of interest.

4. Visual Structure

Top News Headline words by Year view follows a consistently formatted visual structure, with an explanatory block of text at the top of the page, followed by an bubble chart showcasing the relative counts of the words. A bar chart with tooltips helps compare the counts of words in these news headlines. Both the bubble chart and bar chart follow consistent color coding to help user co-relate words. Gradual unfolding of the counts and bubbles helps the user draw a perspective of most used word to the least used word. A filter by Person, Location and Organization helps narrow the keywords and reduce noise. Background colors in these filter choices help drive attention.

Navigation on the left pane and annotations on the top directing the user to follow this navigation help drive further interaction.

Comparison of word counts across Years view is a means to draw meaningful information about trends. Geopolitical news articles such as Russian sanctions and their peak in 2017 and 2018 are annotated in this view to encourage and provide examples to help user draw further trends and conclusions. Legends and tooltips are provided to help navigate with words and their counts.

5. Scenes and Visual Ordering

There are 6 scenes of the narrative visualization. Each of them navigable using a left navigation pane:

- Top Words by Year: This view provides the user with top words used during a particular selected year. A bubble chart showcasing the relative counts of the words. A bar chart with tooltips helps compare the counts of words in these news headlines. Consistent color coding scheme helps tie these charts together. A scene for each of the following years is provided:
 - Year 2021
 - Year 2020
 - Year 2019
 - Year 2018
 - Year 2017
- Compare YoY: A bar chart with annotations that compares the usage of words in News headlines with their counts stacked by year. Year wise color coding helps draw trends. Example of a trend to draw within a stacked bar such as the word “Vaccine” and its counts show a low to high rollout 2020 followed by 2021. Example of a trend across stacked bars would be “Covid”, “Pandemic” showing higher stacks for year 2020 and 2021 and no values for 2017 and 2018 explaining the advent of Covid-19 in the year ensuing 2020.

The user is free to select a given scene at any time using the buttons on the left pane which are always accessible during this narration.

6. Annotations

Various annotations are available over the scenes including the following

- Annotations on the top of the page providing a summary of counts per year. These counts provide the user a context of the data size.
- Annotations on the “Compare YoY” scenes depicting trends on how words and their counts stacked YoY.

All annotations follow the same template and format, to maintain visual consistency and reduce any feeling of disorientation in the scene transitions.

7. Parameters

This narration provides various parameters that the user can choose from. The top word count by Year view provides parameters such as the following that help the users filter and drive the narrative based on specific word categories and hence interests

- Places
- Organization
- Person names

A consistent color coding keeps the user perspective to coloring scheme tied to a particular year and hence tied to changes in state via the above parameters.

8. Triggers

Triggers used in this visualization include the following and each drives a animation to present the scene to show a gradual transition from scene to scene

- Left navigation pane Buttons which allow the user to change to a different scene.
- Filter radio buttons that help user filter words by types such as Place, Name or Organization.
- Tooltips which show data values when the user hovers over individual datapoints