

# The Neutralizer



*Knowledge and the Web, 25/1/2019*

**Daudier Chloé**

**DeHaven Robbert**

**Luqman Elias**

**Tran Tuan-Anh**

# TABLE OF CONTENTS

1. Introduction
2. Method
3. Architecture
4. Data
5. Reference mapper
6. Anaphora-coreference resolution
7. Summarizer
8. ClaimBuster filter
9. Topical similarity
10. Grouping similar sentences
11. Results
12. Conclusion
13. Further work

# 1- Introduction

## ➡ Goal of the Neutralizer

- To educate the user on a given news subject

## ➡ How to achieve this?

- Show the references for the news articles, visually
  - Display relationship between articles and their references
  - May identify common sources between different articles
  - A method to find more articles in the same news subject
- Provide a spectrum of views from a variety of news sources
  - Generate summaries for different news articles
    - General summaries
    - Factcheck-worthy summaries
  - Highlight topically similar sentences in different summaries
    - Do sentences agree with or contradict each other?
  - Display results in a UI-friendly format

## 2- Method

### ➡ Collect data

- Select news subject and identify articles from a variety of sources

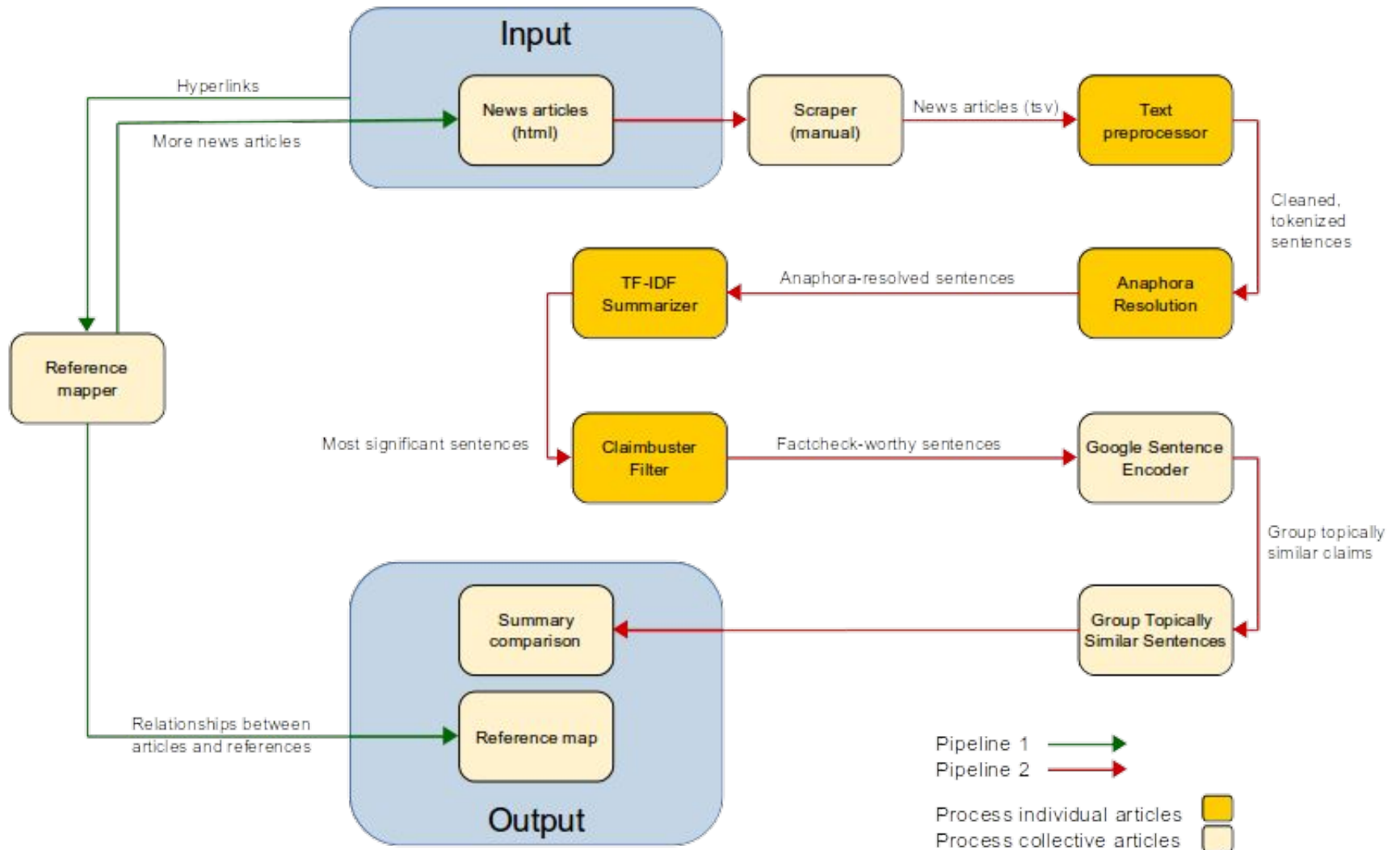
### ➡ Pipeline 1: reference mapper

- Extract urls (i.e. references) from each news article
- Display graph of relationship between articles and their references

### ➡ Pipeline 2: summary comparison

- Extract text, title, url etc. from each article
- Summarize each article
- Run summaries through ClaimBuster filter to generate factcheck-worthy summaries
- Create topical similarity matrix for all sentences in a news subject
- Create topical groupings

# 3- Architecture



# 4- Data

## Data sources

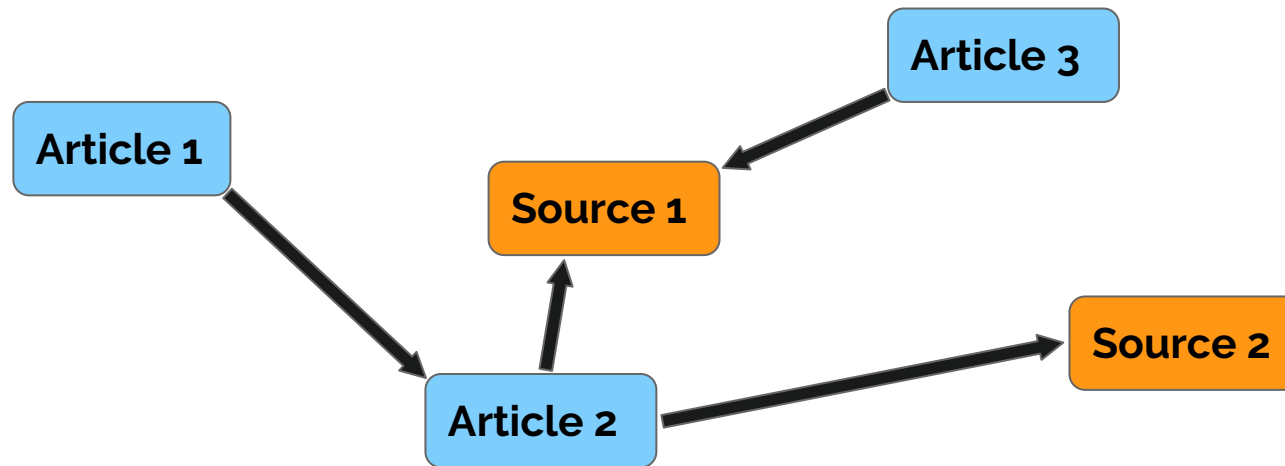
- From mostly US news sites
- Handpicked  $\pm 10$  news articles in 3 controversial news subjects
- News sources:
  - Left: BuzzFeed News, Vox, Huffington Post, Slate, Daily Kos, The Washington Post etc.
  - Center: NPR, BBC News, New York Times, ABC News, The Hill, Bloomberg etc.
  - Right: New York Post, Breitbart, Fox News etc.

## Data scraping

- Manual scraping

# 4- Reference mapper

## ➡ What is reference mapping?



## ➡ Why is this mapping useful?

- Easily view the **most cited sources**
- Show **clusters** that cite the same source (and possibly have the same view on the news subject)
- Indicate possible **lack of sources**
- Help **identify new articles** we can potentially add to dataset

## 4- Reference mapper

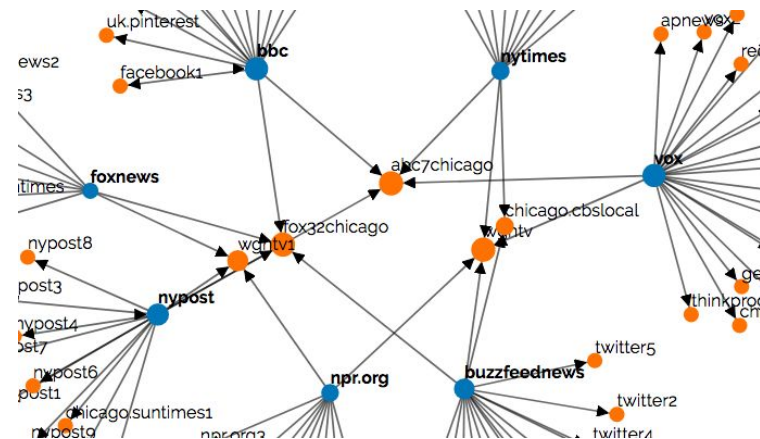
## ➡ Implementation details

1. **Select** ±10 news articles for the given subject from a range of sources
2. **Extract** and **filter** urls from the html of each article
3. Create dictionaries for **nodes** and **links**
4. Display **graph**



```
url1 = 'https://www.buzzfeednews.com/2018/01/11/fox-news-
url2 = 'https://www.foxnews.com/2018/01/11/fox-news-
url3 = 'https://www.vox.com/identical/2018/01/11/fox-news-
url4 = 'https://www.theguardian.com/us-news/2018/jan/11/fox-
url5 = 'https://www.npr.org/2018/01/11/590111111/fox-news-
url6 = 'https://www.bbc.com/news/health-55011111
url7 = 'https://www.nytimes.com/2018/01/11/health/fox-news-
url8 = 'https://www.dailymail.co.uk/health/article-5501111
url9 = 'https://nypost.com/2018/01/11/fox-news-anchors-
```

```
graph = {
    nodes : [
        { ... },
        { ... }],
    Links : [
        { ... },
        { ... }]
}
```

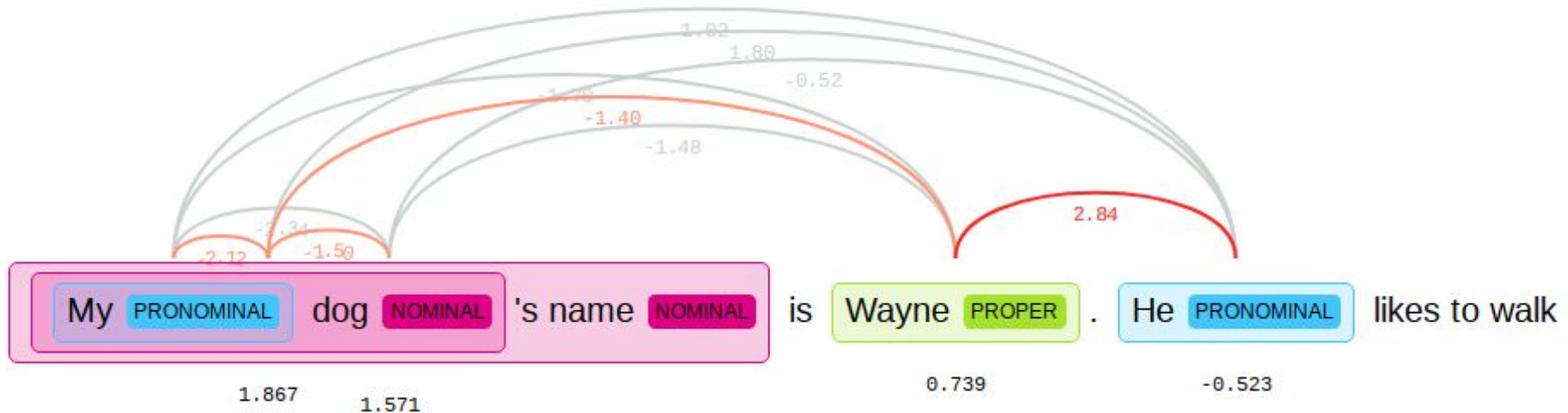




# 5- Anaphora-coreference resolution

## ➔ Resolve anaphora

- Improves readability in summaries
- Example:



## ➔ Implementation details

- Neuralcoref library by Huggingface
- Neural network model

# 6- Summarizer: option 1 (TF-IDF)

## ➔ TF-IDF (term frequency - inverse document frequency)

- Popular term weighting scheme
- Measures how relevant a term is to a particular document
- Background corpus: NLTK Reuters corpus (subset)
- Sentence score: sum of TF-IDF scores for every word

## ➔ Optimizations

- Stopword-filtering
- Only nouns and verbs scored
- Sentences related to title given a score bonus
- Other possible tweaks:
  - Stem/lemmatize terms
  - Apply positional weights

## ➔ Implementation details

- Generate top-5 and top-10 sentence summaries
- Sentence ordering: document order
- Scikit-learn (TF-IDF)
- NLTK (stopword-filtering, POS tagging, Reuters corpus)

# 6- Summarizer: option 2 (TextRank)

## ➔ PageRank

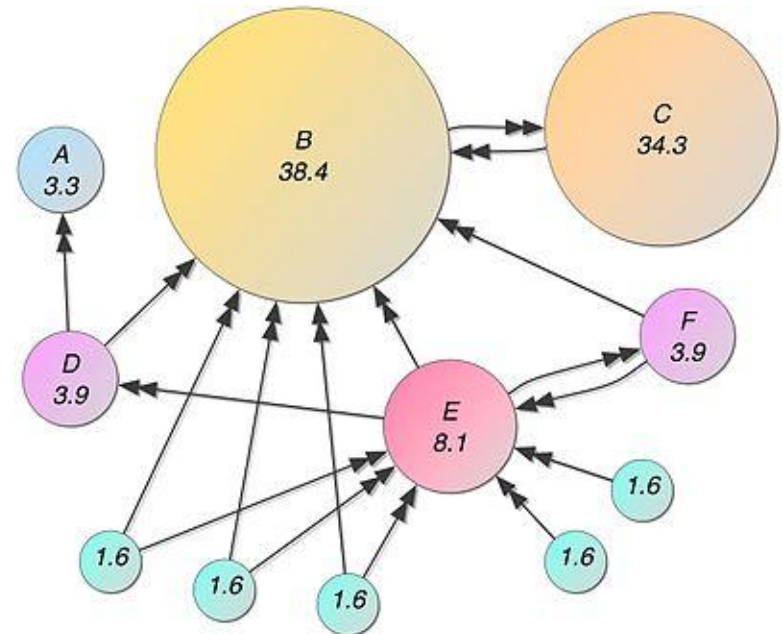
- Measuring the importance of website pages

## ➔ TextRank

- Consider sentences the equivalent of web pages
- The probability of going from sentence A to sentence B is equal to the similarity of the 2 sentences
- Apply the PageRank algorithm over this sentence graph

## ➔ Implementation details

- Pytextrank library



# 6- TF-IDF and TextRank comparison

## ➔ TF-IDF summary

1. An armed security guard at a bar in suburban Chicago was killed by police as he detained a suspected gunman, according to officials and witnesses.
2. After gunfire erupted around 04:00 local time on Sunday, Jemel Roberson, 26, chased down an attacker and knelt on his back until police arrived.
3. Moments after police came on the scene, an officer opened fire on Roberson, who was black, killing him.
4. Sophia Ansari, a spokeswoman for the Cook County Sheriff's office, said police were called to the scene after a fight broke out in the bar and four people were shot.
5. "How in the world does the security guard get shot by police?"

## ➔ TextRank summary

1. An armed security guard at a bar in suburban Chicago was killed by police as he detained a suspected gunman, according to officials and witnesses.
2. Moments after police came on the scene, an officer opened fire on Roberson, who was black, killing him.
3. Friends say Roberson was a musician who had dreams of joining the police.
4. "Just waiting on the police to get there."
5. Illinois State Police's Public Integrity Taskforce have been asked to investigate the shooting, Ms Ansari told the BBC.

# 7- ClaimBuster filter

## ➔ Factcheck-worthy filter

- ClaimBuster scores sentences by how factcheck-worthy they are
- Scoring is done on a scale of 0.0 to 1.0 (the higher the score, the more likely the sentence is to contain factcheck-worthy claims)
- Scores are not context sensitive

## ➔ Examples of sentences that are not factcheck-worthy

- “Here is what we know so far about the attack, the suspect and the victims.”
- “Reports said Cecil Rosenthal liked to greet people at the door of the synagogue before services.”
- “Jonathan Greenblatt, chief executive officer of the Anti-Defamation League, said the group believes Saturday's attack was the deadliest on the Jewish community in US history.”
- “As news of the recount broke, Trump weighed in on Twitter.”

## ➔ Implementation details

- Queried ClaimBuster API
- Threshold of 0.3 used

# 8- Topical similarity: option 1 (SIF)

## ➔ SIF (Smooth Inverse Frequency)

- Based on word embeddings
- Weighted average of word embeddings used
  - Word weight =  $a/(a + p(w))$   
(where  $a$  is a parameter that is set to 0.001 and  $p(w)$  is the frequency of the word from a reference corpus)
- Removes bias from words that are semantically irrelevant

## ➔ Implementation details

- Gensim (Word2Vec embeddings)
- Scikit-learn (cosine similarity)
- NLTK (stopword-filtering)

# 8- Topical similarity: option 2 (Google USE)

## ➔ Google USE (Universal Sentence Encoder)

- Pre-trained deep averaging network (DAN) model
- Can used for semantic similarity and text classification tasks
- Optimized for “greater-than-word length text, such as sentences, phrases, or short paragraphs”
- Advantages over SIF:
  - Supervised model
  - Takes word order into account
  - Models embeddings beyond the level of individual words

## ➔ Implementation details

- Google USE model (1GB)
- Tensorflow

# 8- SIF and Google USE comparison

Sentence 1	Sentence 2	SIF similarity score	USE similarity score
Roberson played music at the Purposed Church in Chicago for the past several years, according to Pastor LeAundre Hill, who tweeted that Roberson "had just played for my grandma's funeral Friday and now he's gone."	Roberson was working to "enough money together for a deposit on a new apartment", said Hunter, the great uncle of Laquan McDonald, a black teenager fatally shot in 2014 by a white Chicago police officer.	0.26	0.43
Four other people were shot and wounded during the incident, including a man who police believe fired a gun before police arrived, the Cook county sheriff's spokeswoman, Sophia Ansari, said.	Four others were shot and wounded, including a man who police believe fired a gun before police arrived, Cook County sheriff's spokeswoman Sophia Ansari said.	0.89	0.98
A police officer fatally shot an armed security guard who witnesses say was trying to detain a man following a shooting at a suburban Chicago bar, authorities said.	An armed security guard at a bar in suburban Chicago was killed by police as he detained a suspected gunman, according to officials and witnesses.	0.58	0.94



# 8- Semantic similarity vs stance similarity

Sentence 1	Sentence 2	Semantic similarity score
Let's go home	Let's go home	1.00
Should we get going	Should we go	0.81
I like strawberries	I do like strawberries	0.97
<b>I do like strawberries</b>	<b>I do not like strawberries</b>	<b>0.95</b>
I like strawberries	Let's have strawberries	0.83

## Semantic similarity --> topical similarity

- Original goal to implement stance detection using semantic similarity methods
- Not sufficient, so adapted semantic similarity methods to measure topical similarity

# 9- Grouping similar sentences

## Manual selection

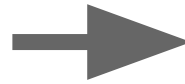
Sentence 1	Sentence 2	Semantic similarity score
A	B	1.00
A	C	0.81
A	D	0.97
B	C	0.95
B	D	0.83

Select sentences with similarity scores of at least  
**THRESHOLD = 0.9**

# 9- Grouping similar sentences

Similarity matrix

A	1				
B	0.84	1			
C	0.53	0.63	1		
D	0.28	0.37	0.60	1	
E	0.23	0.35	0.70	0.69	1
	A	B	C	D	E



Distance matrix

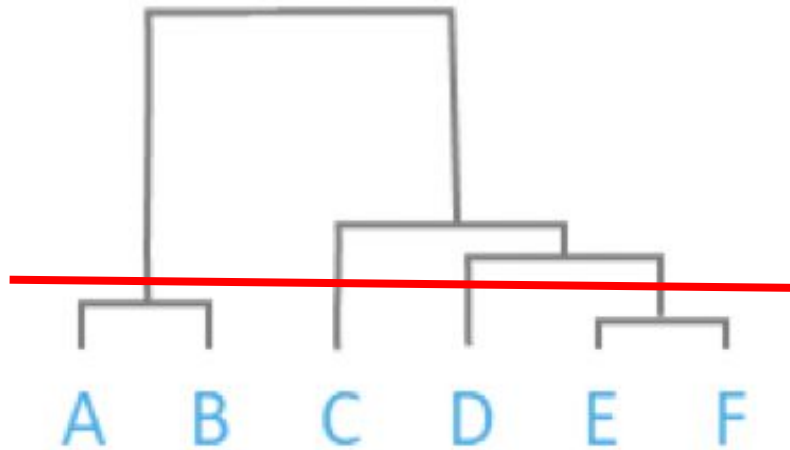
A	0				
B	0.16	0			
C	0.47	0.37	0		
D	0.72	0.57	0.40	0	
E	0.77	0.65	0.30	0.31	0
	A	B	C	D	E

Clustering methods for distance matrix:

- K-Medoids
- **Hierarchical Clustering**

## 9- Grouping similar sentences

Dendrogram



For example: With  
**THRESHOLD = 0.8**  
(similarity), we can get 2  
groups **(A, B)** and **(E, F)**

# 10- Results

 Demo

# 10- Results

➔ Demo

➔ ClaimBuster filtering

News Subject	Average number of sentences filtered out by ClaimBuster per summary	
	Top-5 summary	Top-10 summary
Accidental shooting	0.33	1.44
Florida midterm elections	0.00	1.11
Synagogue shooting	0.25	1.28

- Significant majority of sentences in that made it into summaries are factcheck-worthy claims

# 11- Conclusions

## ➔ Views fairly consistent between news articles

- Somewhat surprising given that we picked articles from sources across the political spectrum on what we thought were controversial subjects
- Even questionable sources generate accurate news sometimes. How does the user know when they're being misinformed? One of the major problems with fake news?

## ➔ Larger dataset needed for more testing

- May discover problems we didn't find with our limited dataset
- May also better showcase the abilities of the Neutralizer e.g. display conflicting views from different sources

## ➔ Viable proof-of-concept

- User is exposed to a variety of views on a given news subject
- User can see how the articles are related to their references and possibly each other in a visually clear manner

# 12- Further Work

## Data

- Collect larger dataset (more new subjects, more articles per news subject)
- Automate finding news articles
- Automate scraping (*newspaper* library from MIT)
- Use related news articles found by reference mapper to expand dataset

## Perform stance detection

- Identity consistent and conflicting views automatically

## Summarizer tweaks

- Stem/lemmatize terms
- Apply positional weights



Thank you for your  
attention!

Questions?