Algorithmic Analysis of News Headlines by Artificial Intelligence Systems

A Master’s Degree Project Proposal

Chinmay Pai

Interactive Media and Game Development

Worcester Polytechnic Institute

Worcester, MA

Committee

Advisor - Brian Moriarty

Reader – Gillian Smith

Reader – Jacob Whitehill

2019

# Abstract

“Spin” is a form of propaganda that uses biased language to influence public opinion. This thesis will conduct research intended to inform the design of a prospective game which will allow players to explore the dynamics of spin. A Web application will be developed to scrape daily headlines from popular US new sources and submit them to AI systems such as Watson, Google Cloud and Amazon Comprehend for analysis of their emotional bias. The same headlines will be submitted to a demographically diverse population of American readers for similar evaluation. The results of these evaluations will be compared to identify the strengths and weaknesses of the AI analysis, and help determine which forms of bias should be used to improve the clarity and persuasiveness of the game.

# Introduction

The spread of propaganda has significantly affected the Western world in recent years. The vote for Great Britain to leave the European Union, the rise to power of US President Donald Trump and other autocratic leaders, and controversies surrounding climate change and immigration, have all been molded by propaganda, often amplified by the widespread adoption of social media technologies.

However, propaganda is not a new phenomenon, and its use is not limited to the West. Developing countries also contain powerful anti-establishment groups who are widely perceived as being more reliable news sources than the ‘compromised’ mainstream. Having seen several friends and acquaintances in India affected by spin playing into culturally established narratives, it is clear that propaganda is effective at exploiting psychological blind-spots in similar ways around the world.

I am interested in creating a digital game that allows players to manipulate real-time headlines from a variety of real-life news outlets with differing levels of bias. My hope is that players experiencing the game will be able to see how spin spreads, and how its effects might be inhibited.

# Prior Research

Most of the work in the field of the analysis of fake news has been either technical – understanding the spread of fake news through new mass media such as Twitter – or sociological – understanding the nature of fake news and how it influences the people it targets to achieve its ideological aims. Sobieraj and Berry have spoken about how the spread of outrage across all forms of media in the United States, such as television, talk radio, newspaper editorial columns, and political blogs, has been a significant measure of the success of these media in the form of viewership and clicks, which seems to correlate with the generally accepted idea that ‘outrage sells’, and have also measured so-called incivility from both sides of the political spectrum, concluding that while both the left and the right use similar tactics, the right have been shown to use it in greater quantities. Langin has found that fake news spread is not, contrary to popular belief, primarily spread by bots, but by actual humans, whether for malicious or ignorant reasons.

There have also been several ideas regarding the methods by which fake news might be contained. These include the idea of “guardians”, proposed by Vo and Lee, users who can recommend verified facts to users in response to popular misconceptions/fake news about popular figures. Websites such as Politifact and Snopes have taken up this role with mixed success, with allegations of bias from both sides of the political spectrum. Baum also speaks about the possibility of weeding out fake news using algorithmic methods, which utilize bots to either correct or remove fake news, which would require collaboration between sociological academia and computer science experts, as well as careful balancing to ensure that governments do not subvert these systems to tyrannical ends.

# Research Design

The research which is to be carried out is intended to be background research for the design of a serious game. The game under discussion, tentatively titled “The Foghorn”, has been designed as a puzzle game where the player controls the head of a media organization whose aim is to rise from a small-town newspaper to a media empire spanning the domains of radio, television, and new media such as the internet using the power of fake news and information warfare to spread their ideology. The game mechanics involve using the forms of media the player has access to and the news articles of the day to profit off prevailing sentiments and propagate the agenda of choice for the player, with occasional observations regarding how it affects other people, especially the less fortunate.

A screenshot of a cell phone

Description generated with very high confidence

*A basic image description of the gameplay system for newspapers. Individual news items (on the right) can be dragged to the position deemed appropriate, as new stories keep coming in every 30 minutes (10 seconds). The timer above indicates how much time the player has before the paper is sent to print.*

As part of being a serious game, a large part of how the game seeks to tie the game to its real-life context is the mechanic where the game draws its headlines from the actual, real-world headlines of the day from major news organizations. Analyzing these headlines in order to find their nature in accordance with research data available requires carrying out this kind of research in order to avoid valid accusations of personal bias.

The project aims to scrape the headlines similar to the method suggested by Robin and analyses them in terms of their ability to cause outrage as well as their conformity to journalistic standards. To measure these in a manner in any way compliant with ethical standards, this research must be carried out to study such topics in the context of existing research around them and create the required criteria by which the headlines can be gauged.

# Background Preparation

In preparation for the start of the thesis, I have carried out preliminary research in order to display early data results for several artificial intelligence systems which can be used to create guesses for the results that are obtained when analyzing headlines from several sources, such as Google Cloud, IBM Watson and Amazon Comprehend. The headlines were drawn from several news sources such as CNN, Fox News, The New York Times, The Washington Post, Mother Jones and Breitbart. The raw results are available in the appendix provided.

In most cases, the results of the analysis were surprising – even media sources considered highly partisan typically showed high levels of neutrality in the analysis of their headlines as per Amazon Comprehend, Google Cloud, and IBM Watson. Therefore, simple text analysis of headlines and news articles, while informative, is not sufficient to distinguish news sources in terms of “outrage value”, so to speak. Therefore, other factors which affect their reporting may need to be considered.

An observation of the websites of the news organizations, pictures of which are archived in the appendix provided, shows other potential factors which can be seen on their websites. The primary factors observed are positioning and frequency. Positioning certain headlines at the part of the screen which will immediately attract attention affects the perceived importance of the headline in the view of its readership – for the most obvious example, the story placed at the top of the front page is the headline that the organization wishes to show front and center, while stories below may be considered of lesser value no matter what their actual nature is. The same story (for example, the December 8, 2018 announcement that John F. Kelly would step down from his post as White House Chief of Staff at the end of 2018) was positioned at different positions on different news websites – The New York Times and The Washington Post positioned it front and center, the Wall Street Journal placed it at the top but giving equal importance to other headlines, Fox News placed it as a secondary headline which required scrolling to view underneath the main story of James Comey’s testimony to the House Oversight and Judiciary Committee, while Breitbart placed it underneath an article speaking about how Bill and Hillary Clinton had to sell tickets to their speaking engagements through Groupon.

# Evaluation

The essential question which must be answered by this evaluation is:

Is the algorithm created by the research measuring the characteristics of the news headlines provided, and is the conclusion of the algorithm shared by as large and demographically diverse a group of Americans?

Since the analysis of the news headlines is, to a large extent, subjective, I will require human evaluation to be carried out in order to evaluate the data extracted from the analysis of the headlines. I do not believe it will be sufficient to only poll the students of Worcester Polytechnic Institute, as there is highly likely to be significant bias in the results of polling a highly specific demographic (largely 18-30 students in Massachusetts). Therefore, in order to get greater demographic variety, I intend to spread my survey into the internet using the Mechanical Turk system, with filters only allowing for US-based IP addresses to answer the survey. To ensure that the test takers are compensated for their efforts to contribute to my projects, I will ensure that all survey takers will be paid the minimum wage for the state of Massachusetts i.e. $12 per hour. I understand that I will have to consider the possibility of bots interfering with accurate data collection, but I believe that that can be designed around using pictorial representations (as shown in the appendix) which would be very likely to confound bots and is the most practical method by which a large sample size of appropriate variation can be built.

Evaluation will be carried out by showing a sample of headlines and news articles from a wide variety of news publications of different ideological bent, and asking the survey-taker to rate the headlines on different parameters, such as:

1. Outrageousness.
2. “Clickbait” (the perception that the headline has been written in such a way so as to entice the user to click on it)
3. Factual accuracy.
4. Bias.
5. The level to which it attempts to push its agenda.

The comparison of the perception of the sample group and the analysis of the algorithm on the same headlines and news articles would be part of the observations and inferences on the final report of the thesis.

# Timetable

|  |  |
| --- | --- |
| Event | Deadline |
|  |  |
|  |  |
| Final Thesis Presentation |  |
| Final Thesis Submission |  |

# Works Cited

Sarah Sobieraj & Jeffrey M. Berry (2011) From Incivility to Outrage: Political Discourse in Blogs, Talk Radio, and Cable News, Political Communication, 28:1, 19-41, DOI: 10.1080/10584609.2010.542360

Baum, M., Greenhill, K., Metzger, M., Pennycook, G., Rothschild, D., Schudson, M., Sloman, S., et al. (n.d.). The science of fake news. Science, 359(6380), 1094–1096. doi:10.1126/science.aao2998

N. Vo, and K. Lee. The Rise of Guardians: Fact-checking URL Recommendation to Combat Fake News. SIGIR, July 2018

Rose-Stockwell, Tobias (2017) This is How Your Fear and Outrage Are Being Sold For Profit, Medium - https://medium.com/@tobiasrose/the-enemy-in-our-feeds-e86511488de

Langin, K. (n.d.). Fake news spreads faster than true news on Twitter—thanks to people, not bots. Science. doi:10.1126/science.aat5350

Robin, Alexander (2017) How I Scraped 7000+ Articles from a Newspaper Website Using Node, FreeCodeCamp - <https://medium.freecodecamp.org/how-i-scraped-7000-articles-from-a-newspaper-website-using-node-1309133a5070>

# Resources for Future Research

Jang, S., & Kim, J. (n.d.). Third person effects of fake news: Fake news regulation and media literacy interventions. Computers in Human Behavior, 80, 295–302. doi:10.1016/j.chb.2017.11.034

Mihailidis, P., Viotty, S., & Payne, J. (n.d.). Spreadable Spectacle in Digital Culture: Civic Expression, Fake News, and the Role of Media Literacies in “Post-Fact” Society. American Behavioral Scientist, 61(4), 441–454. doi:10.1177/0002764217701217

Vargo, C., Guo, L., & Amazeen, M. (n.d.). The agenda-setting power of fake news: A big data analysis of the online media landscape from 2014 to 2016. New Media & Society, 20(5), 2028–2049. doi:10.1177/1461444817712086

Brigida, M., & Pratt, W. (n.d.). Fake news. North American Journal of Economics and Finance, 42, 564–573. doi:10.1016/j.najef.2017.08.012

M. Rajdev, and K. Lee. Fake and Spam Messages: Detecting Misinformation during Natural Disasters on Social Media (short paper). WI. December 2015.

# Appendix 1: Semantic Analysis of News Headlines

The following headlines were semantically analysed by IBM Watson, Amazon Comprehend, and Google Cloud Services. Both the headlines themselves as well as the entire article were analysed where possible, but at minimum the headlines were analysed.

The headlines selected as samples are as follows:

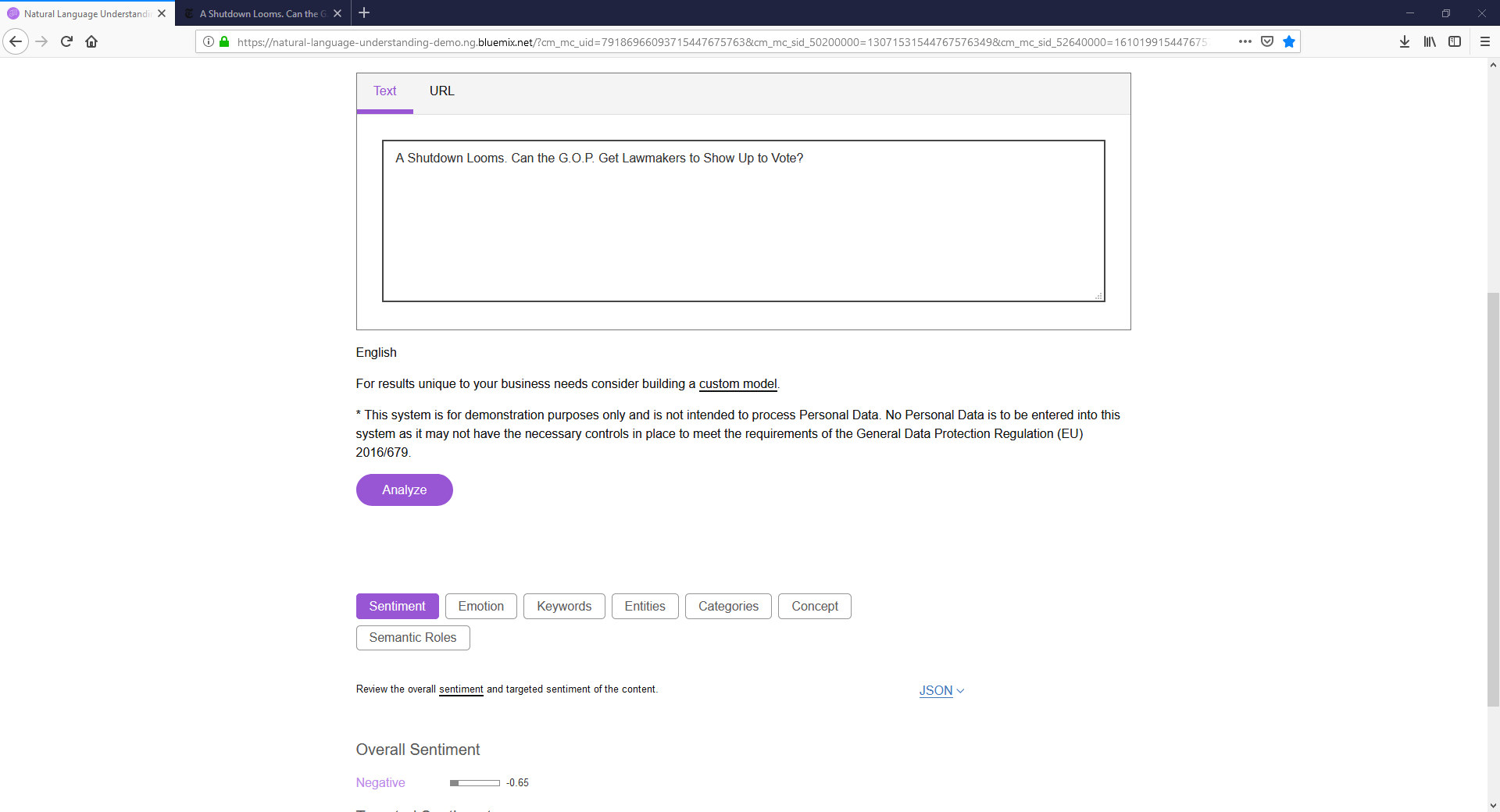
Hirschfeld Davis, Julie and Cochrane, Emily; “*A Shutdown Looms. Can the G.O.P. Get Lawmakers to Show Up to Vote?”*, The New York Times, December 16, 2018: <https://www.nytimes.com/2018/12/16/us/politics/congress-trump-shutdown.html?action=click&module=Top%20Stories&pgtype=Homepage>. This headline is hereafter referred to as **NYT**.

McCarthy, Tyler; Donald Trump Tweets NBC, ‘SNL’ should be tested by courts after Christmas parody sketch, Fox News, December 16, 2018: <https://www.foxnews.com/entertainment/donald-trump-tweets-nbc-snl-should-be-tested-by-courts-after-christmas-parody-sketch>. This article is hereafter referred to as **FOX.**

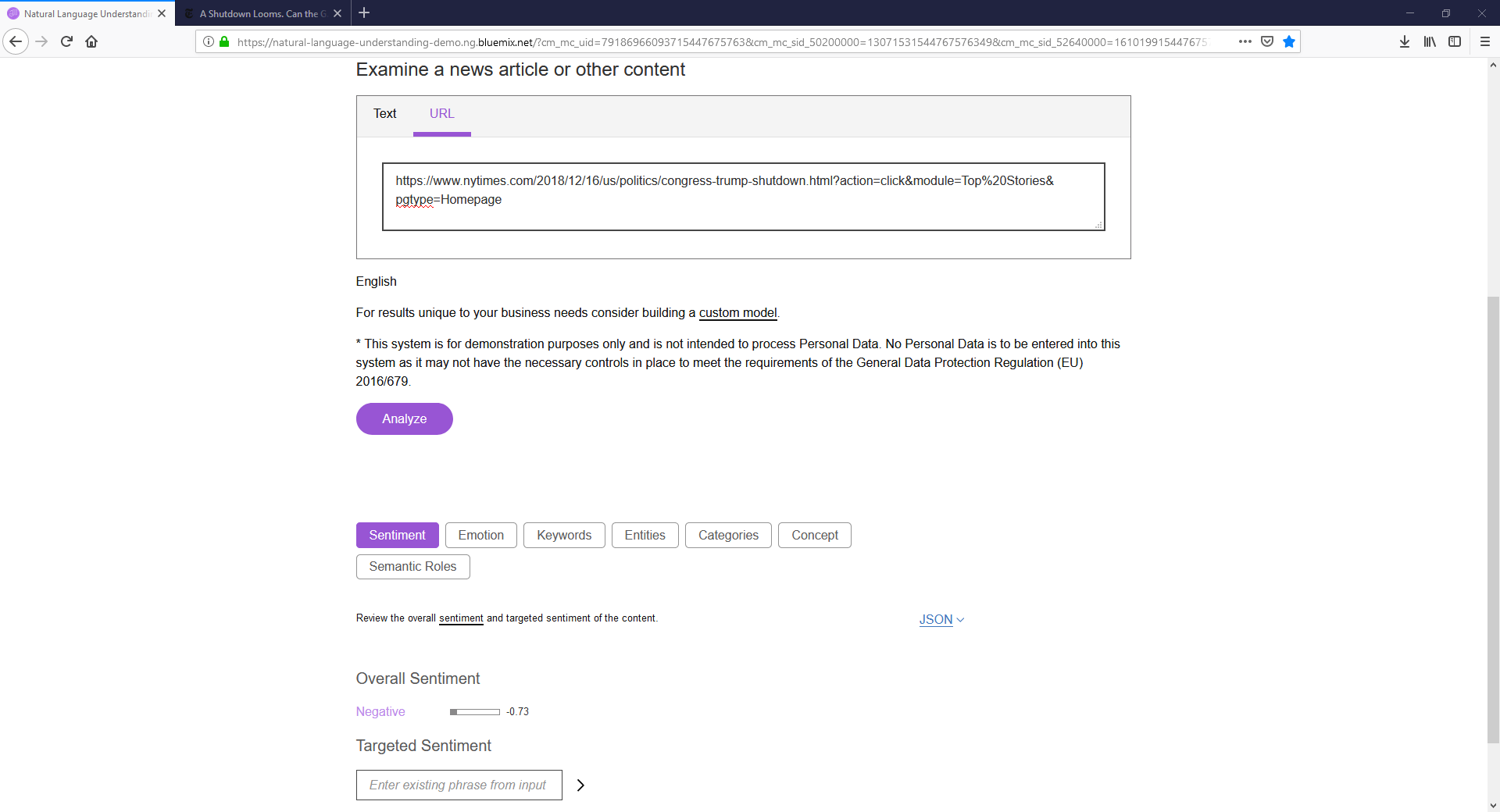
The results are as below:

## IBM Watson

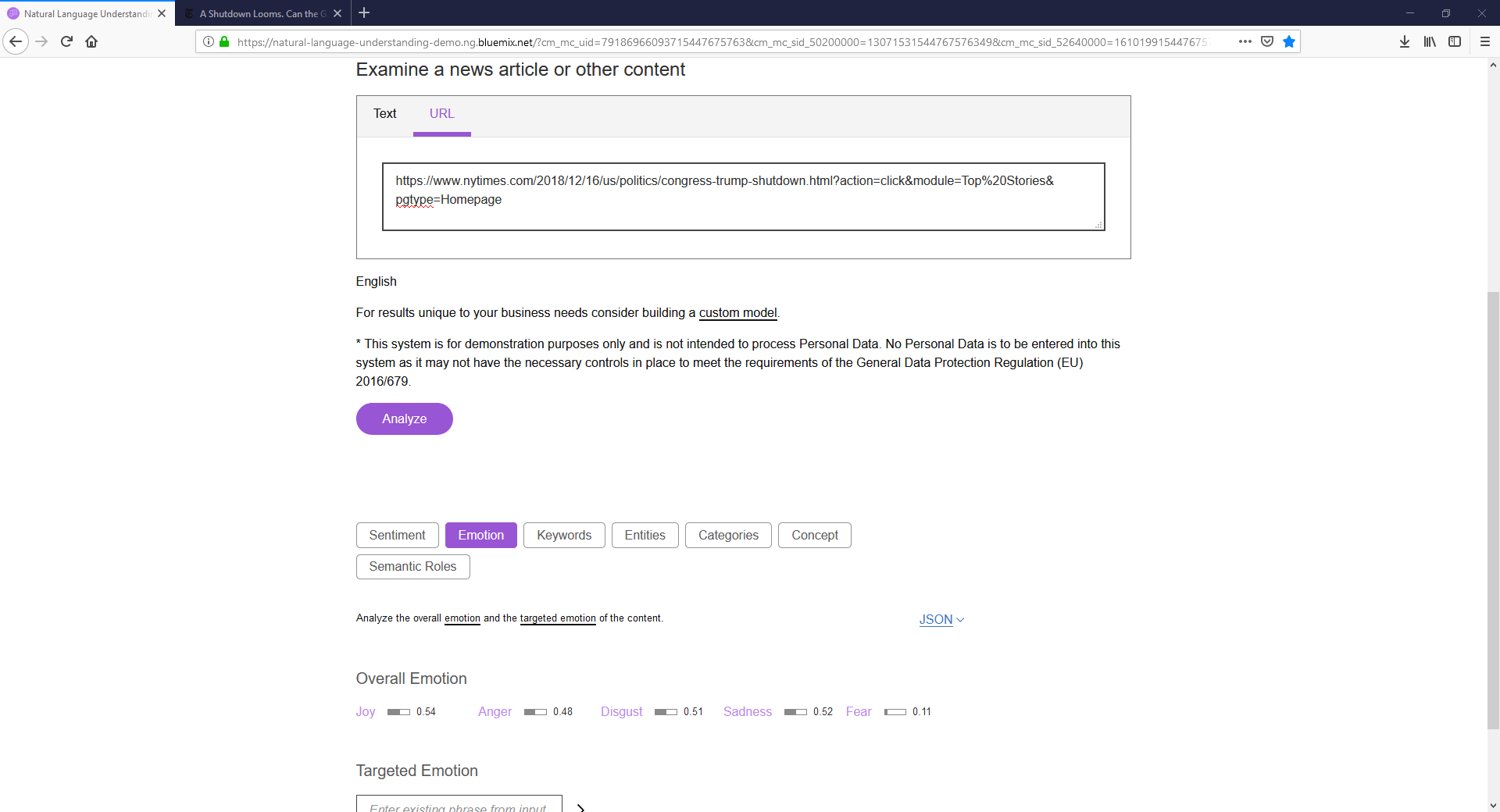
### NYT, Headline Sentiment Analysis



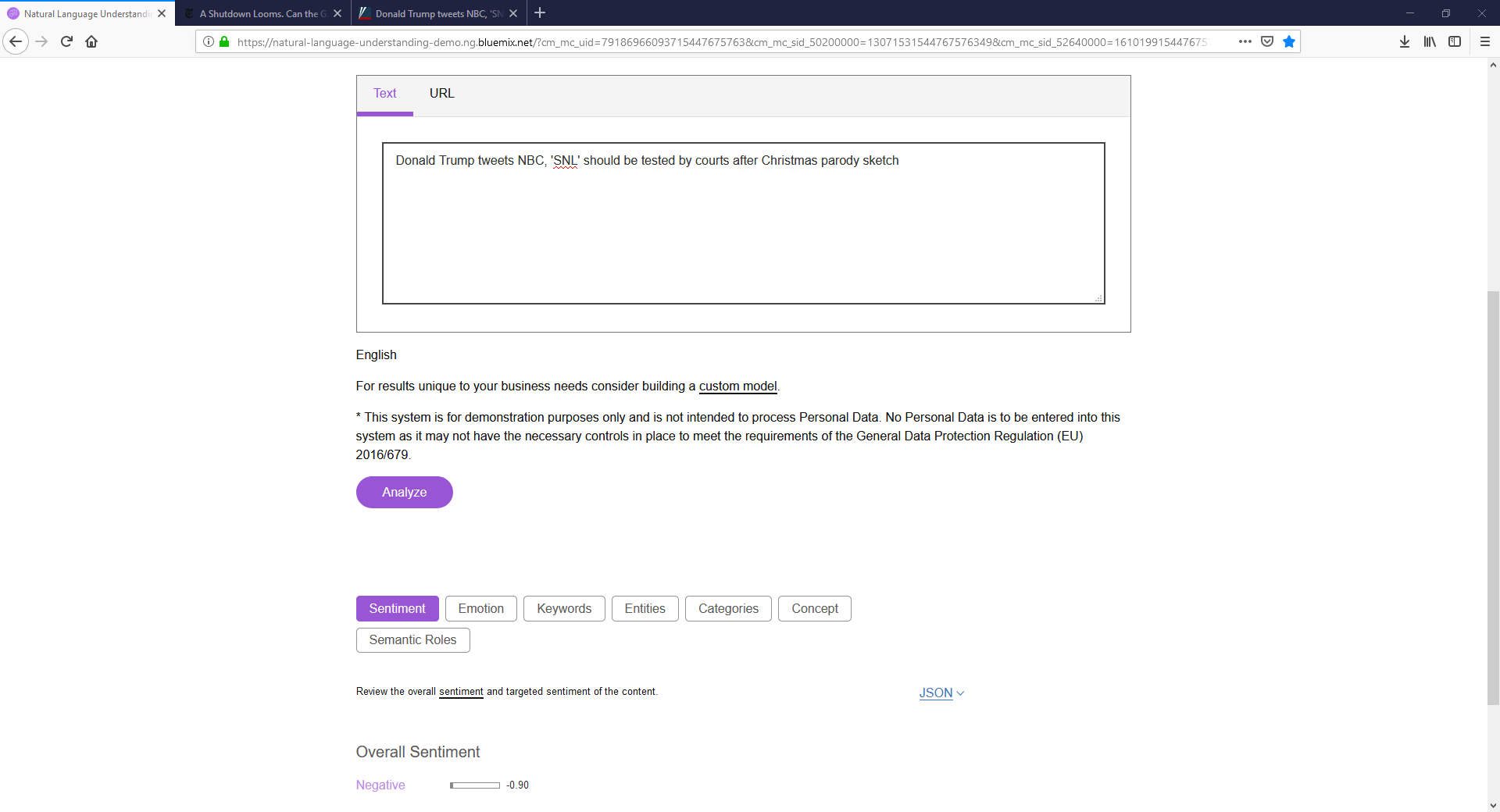
### NYT, Article Sentiment Analysis



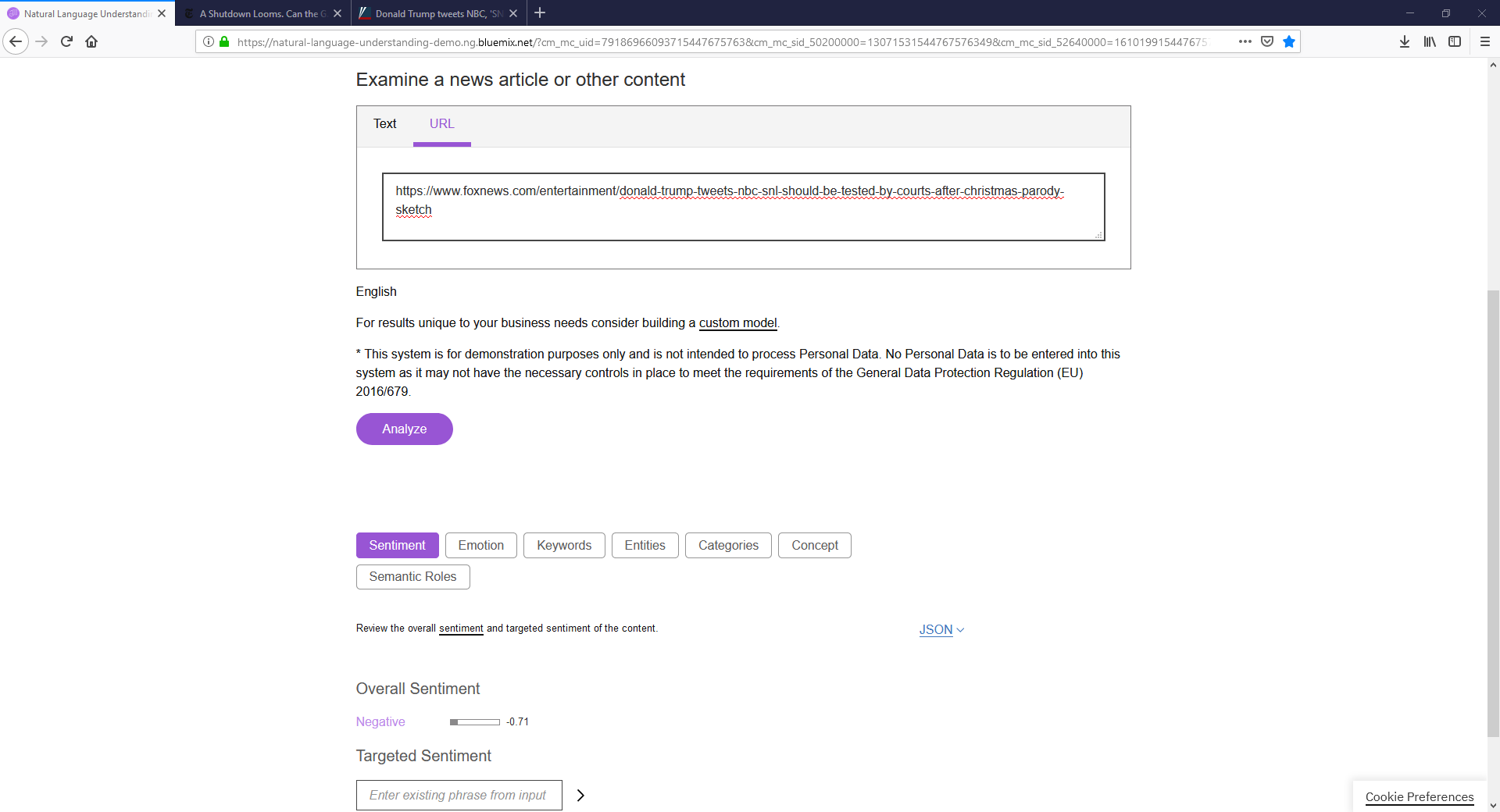
### NYT, Article Emotion Analysis



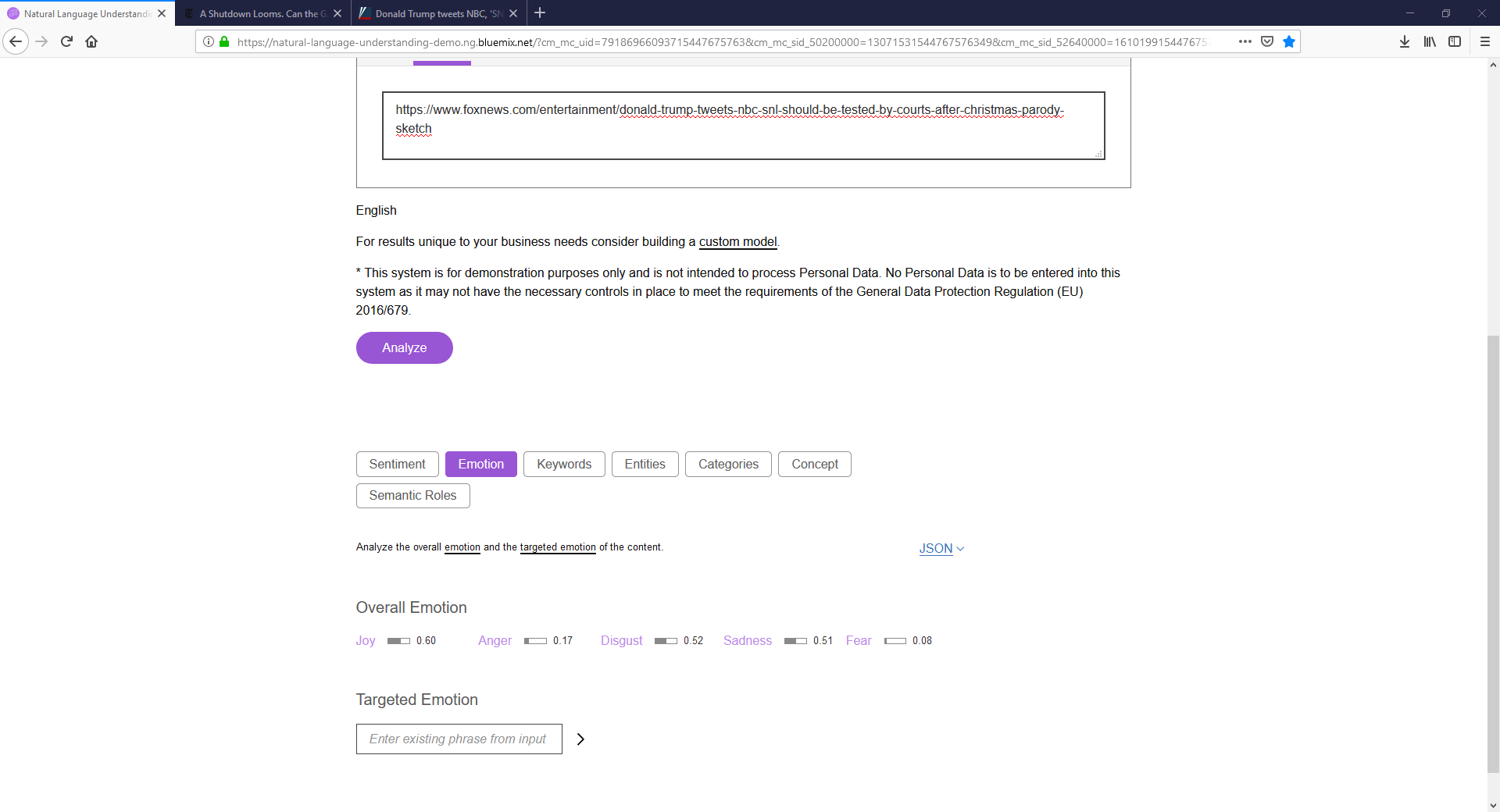
### FOX, Headline Sentiment Analysis



### FOX, Article Sentiment Analysis

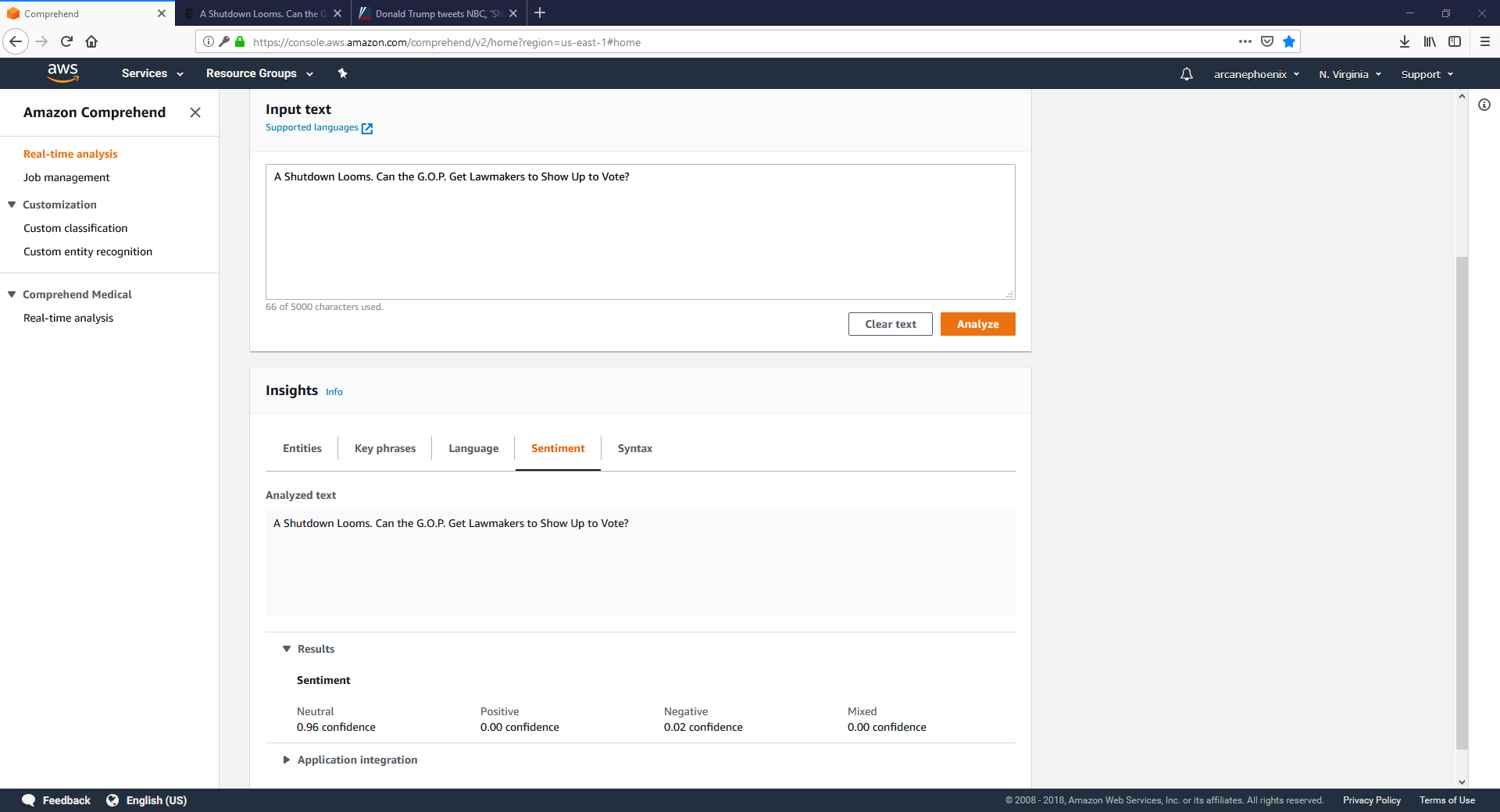


### FOX, Article Emotion Analysis

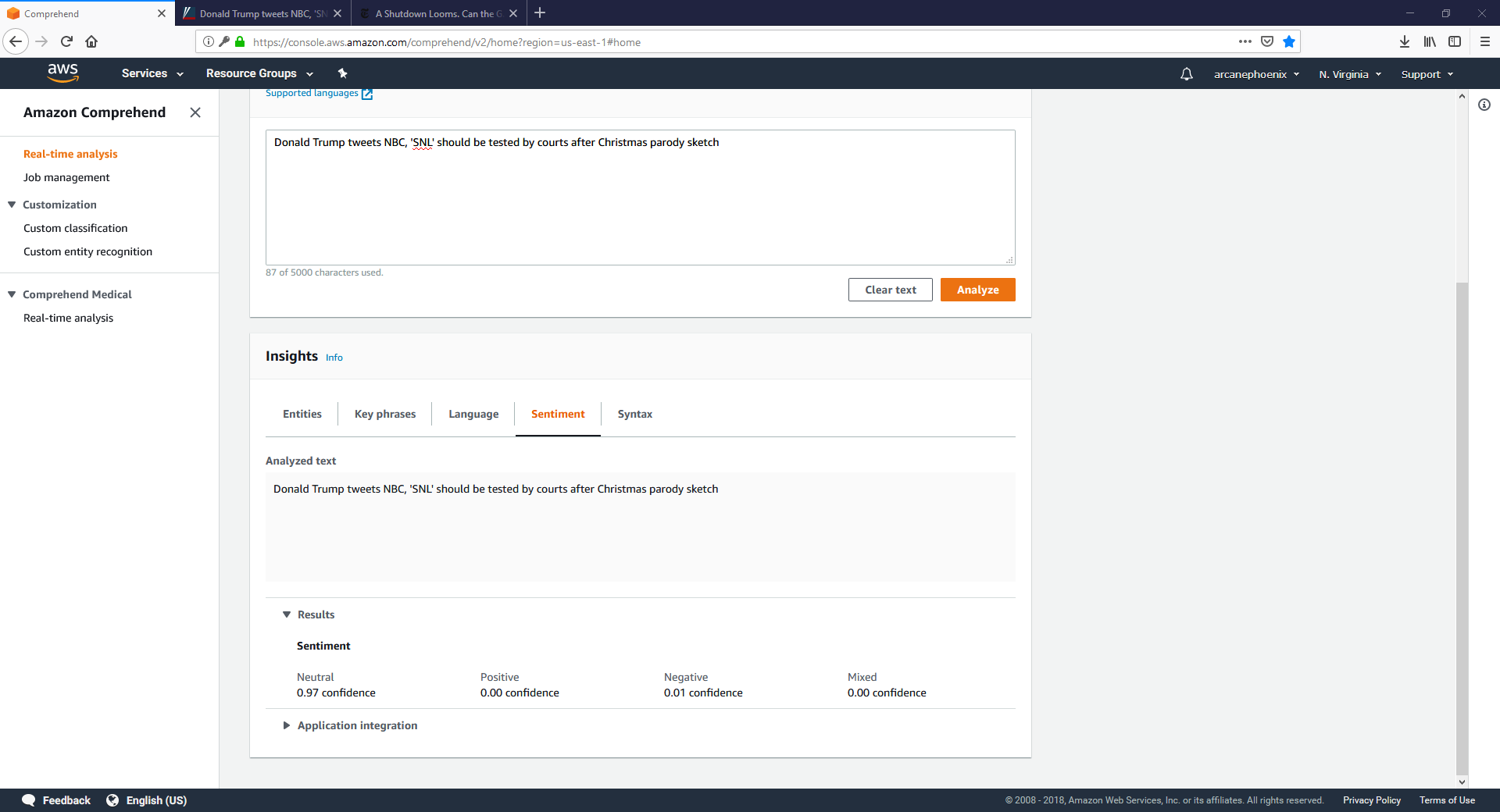


## Amazon Comprehend

### NYT, Headline Sentiment Analysis

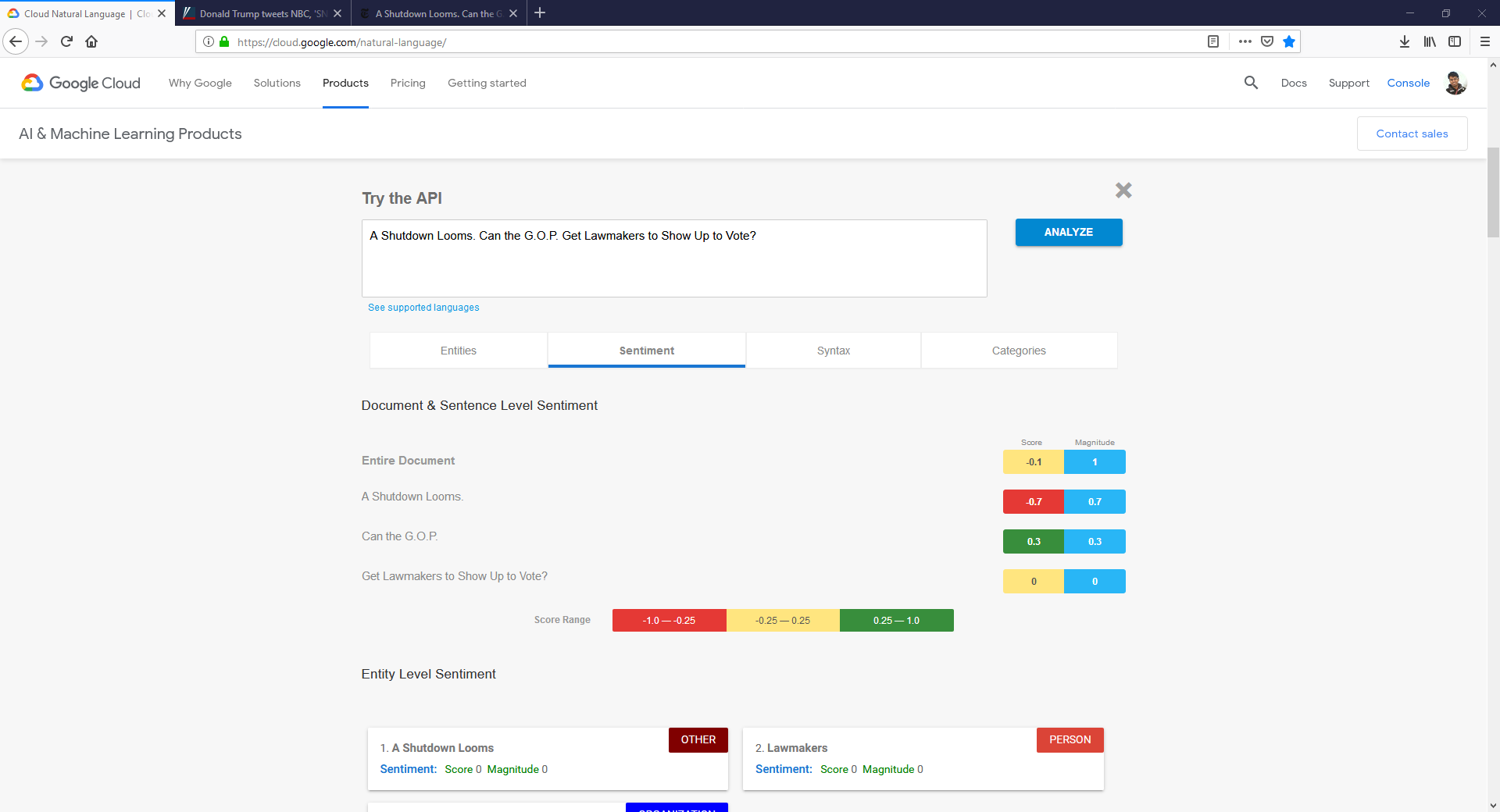


### FOX, Headline Sentiment Analysis

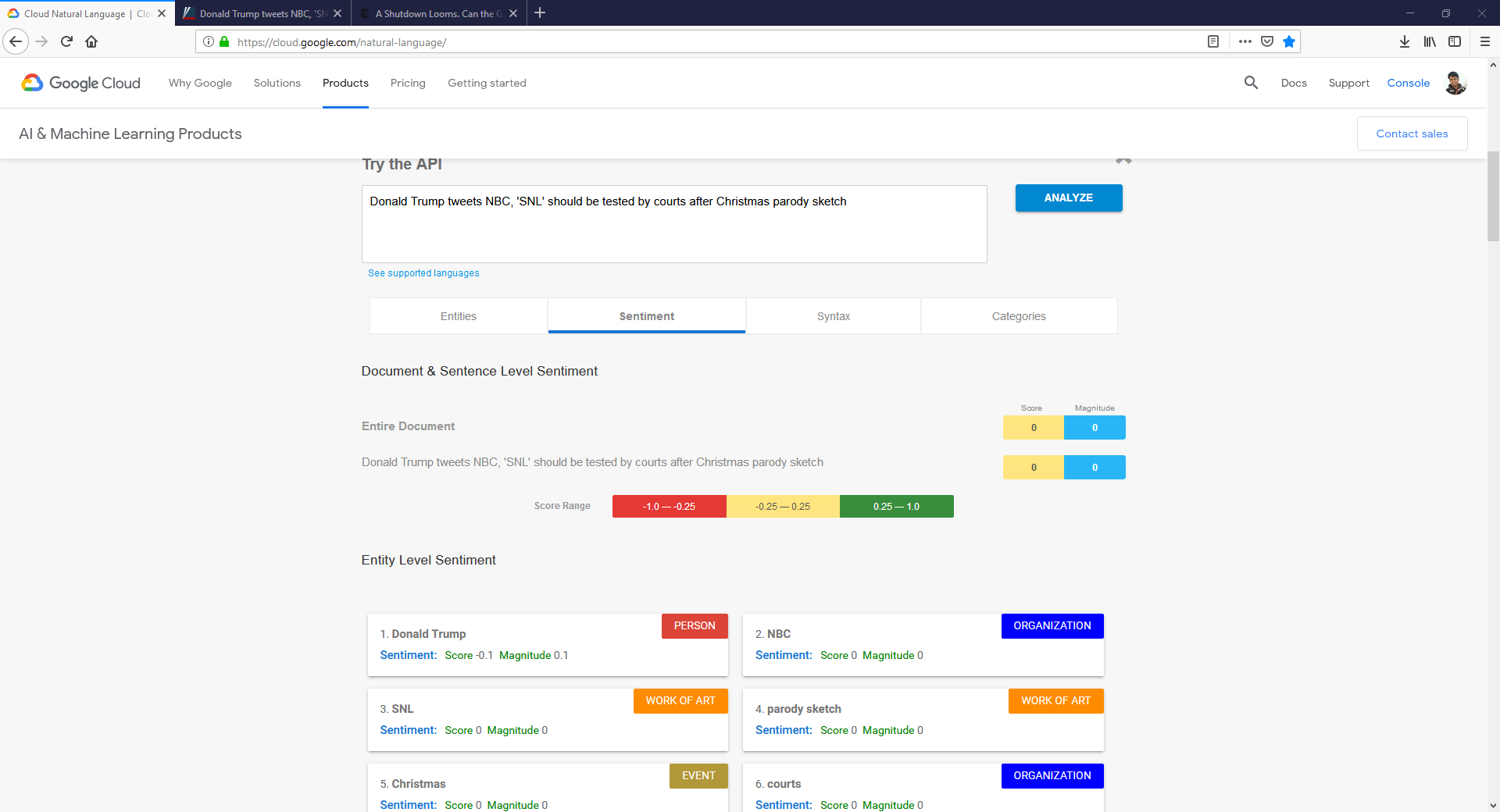


## Google Cloud Services

### NYT, Headline Sentiment Analysis



### FOX, Headline Sentiment Analysis



# Appendix 2: Pictorial Representation of Survey Questions

In order to solve the problem of bots on the Mechanical Turk skewing the data received from the survey evaluation, I will be using pictorial representation to represent the emotions which the survey-taker must interpret in order to answer the question accurately. I believe that this approach will be difficult for bots taking the survey to understand and perceive, and they would prefer to go for simpler alternatives. The same question can also have the solution order changed to further confuse any bots.

The following is a representation of the same question attempting to gauge emotional reaction from the survey taker, but in two different iterations.



