



News for a Green New Deal

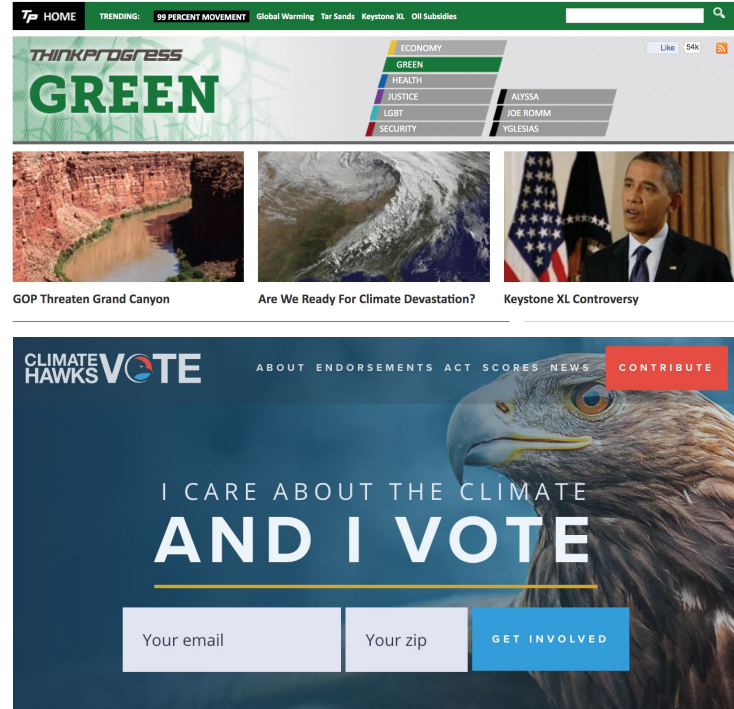
Brad Johnson
@climatebrad
Flatiron School
January 2020



Goal: Find and Organize Climate-Related News

As a long-time climate journalist and activist, I have worked for years to transform the firehose of daily news into coherent stories that provide space for action.

The goal of this project is to organize and classify news articles and potentially other written material by climate-related subtopics based on textual content.



Why is this important?

Man-made global warming is the existential challenge of our time. To build a carbon-free, resilient, and just society, citizens need to learn how the fossil-fuel economy interacts with every aspect of our lives, and how the damages of the pollution are reshaping our world.

Most news and media is not organized to tell that story coherently. For example, extreme weather, political corruption, and energy finance are treated as disjoint, not interconnected topics.

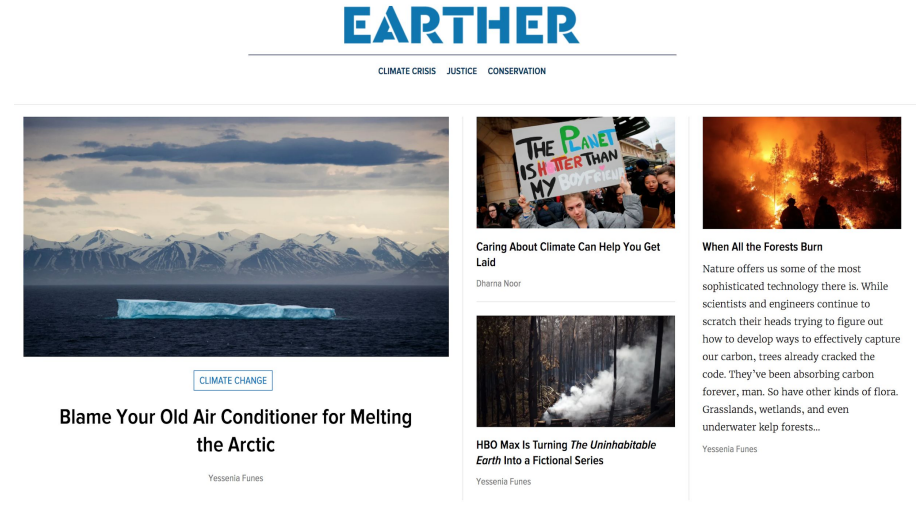
This project is a step towards a high-quality climate news aggregator, applicable to other areas of interest as well.

Steps

- **Scrape** training data: Earther.com articles
- Create training set: K-Means **cluster** on article keywords
- **Train** classifier model on article text
- **Deploy** model for use with any text

Training Set: Earther.com

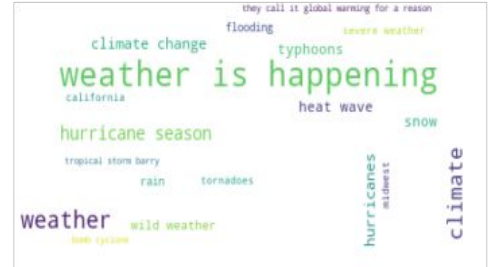
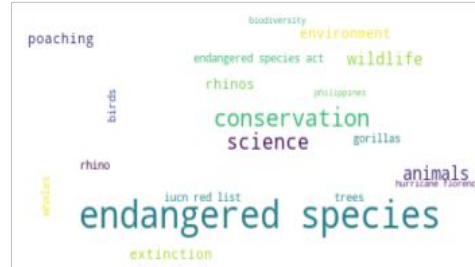
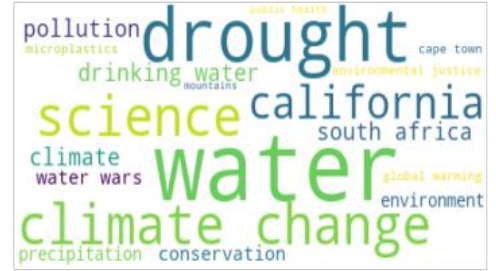
Earther is a website dedicated to a broad range of climate-change related content, from politics and technology to extreme weather and racial justice. There have been over 3,000 articles published between August 2016 and January 2020, each tagged with multiple keywords.



Clustering on Keywords

Earther.com articles have keywords, but there are too many to be useful as topic identifiers on their own. So I analyzed the keywords to find 30 clusters of topics.

Keyword clusters (size reflects frequency)



Topic Classification of Text

Using natural language processing, the model analyzes text and returns probability ratings for associated topics.

Deployed at:

green-new-deal-news.herokuapp.com

GREEN NEW DEAL NEWS CATEGORIZER

Identified topics

- 1 Arctic: 53%
- 2 Meteorology: 21%
- 3 General Climate: 9%

Categorized text

<https://news.miami.edu/stories/2020/01/examining-the-link-between-climate-change-and-extreme-weather.html>



Further Work

- Improve topic generation
- Climate/not climate classifier
- Scrape from Twitter / RSS feeds
- Database-driven website
- Dynamic topic generation
- Highlight most-relevant words in given text
- Categorize images based on topic

Other applications:

- Sort parody from real news
- Develop categories for other subject areas

Questions?

Brad Johnson
@climatebrad
climatebrad@gmail.com
