

Good Press/Bad Press

Renat Khalikov, Patrycja Krawczuk, Sarah Mathew, Ricky Poon

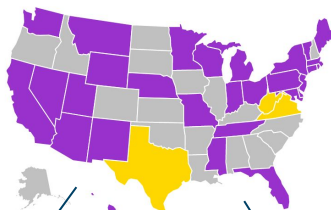




DEMO

OUR JOURNEY

BAD PRESS



Renat & Patrycja



Renat & Patrycja



Natural Language
Analyses with NLTK

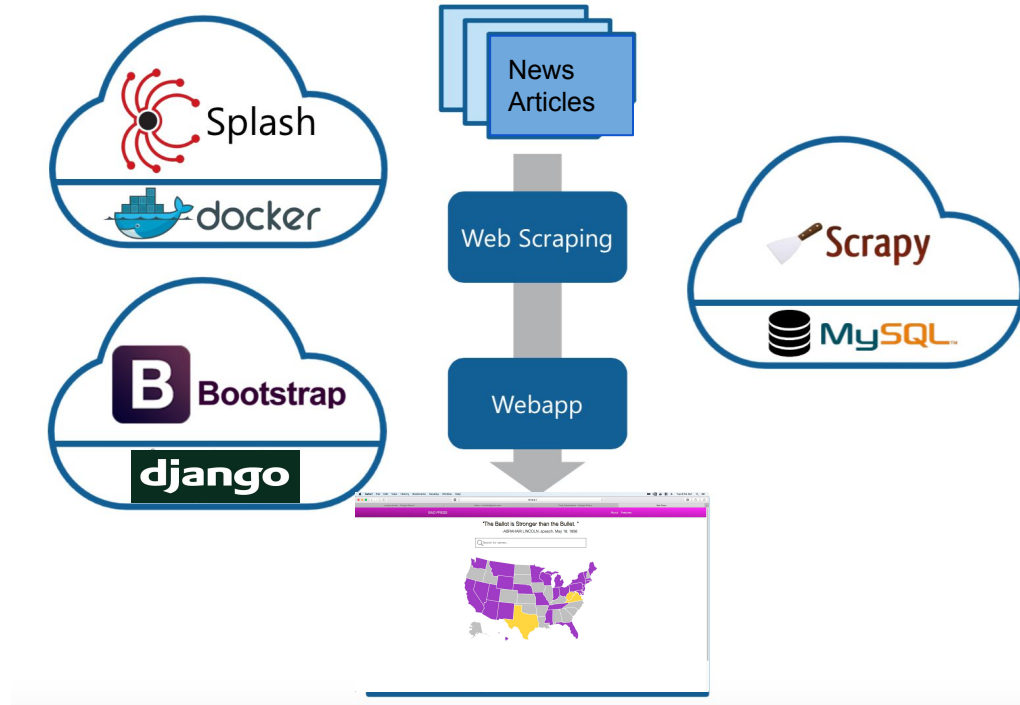
Sarah



Ricky

RENAT:

Scrapy



RENAT:



FOX NEWS FOX BUSINESS FOX NEWS GO FOX NEWS RADIO FOX+NATION FOX NEWS INSIDER LOGIN

FOXNEWSQ http://www.foxnews.com/search-results/search?q=don+blankenship

136 results found for

don blankenship

Advanced Search ▾ Date Relevance

President Trump rejects West Virginia's Don Blankenship

President Trump tweets his West Virginia Senate endorsements. Peter Doocy has the latest developments from primary day.
May 8, 2018

Don Blankenship dismisses Trump comparison to Roy Moore

President Trump urges West Virginia voters not to vote for Blankenship; Peter Doocy reports from Charleston, West Virginia.
May 7, 2018

Trump urges West Virginia to reject Don Blankenship

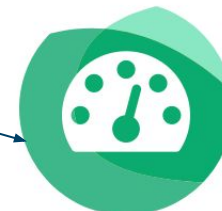
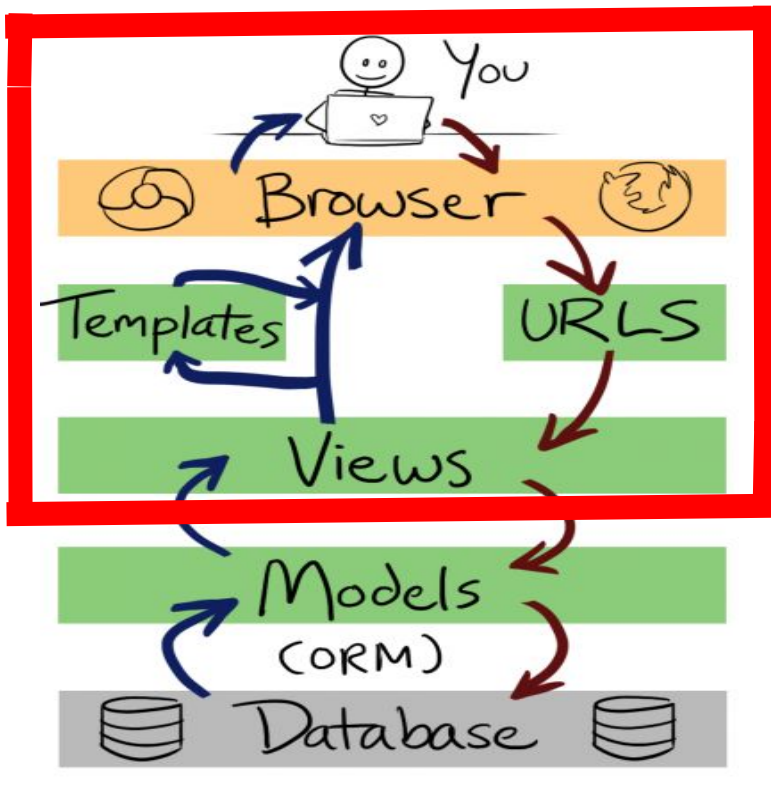
President weighs in on GOP primary race to flip Senator Joe Manchin's seat; Peter Doocy reports from Charleston, West Virginia.
May 7, 2018

1 2 3 4 5 6 7 →

The
New York
Times

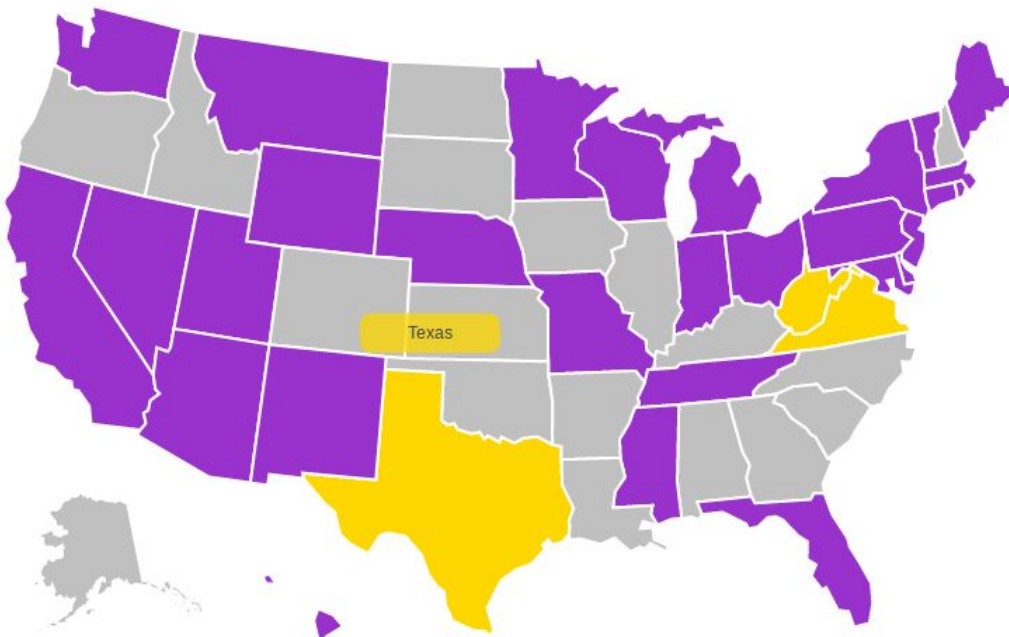
```
<div ng-if="!article.adContent" class="search-article ng-scope">
  <div class="search-info responsive-image" ng-class="article.image ? 'responsive-image' : ''">
    <h3>
      <a ng-bind="article.title" ng-href="http://video.foxnews.com/v/5781845646001/"
        class="ng-binding" href="http://video.foxnews.com/v/5781845646001/">President Trump
        rejects West Virginia's Don Blankenship</a> = $0
    </h3>
    <!-- ngIf: article.description -->
    <p ng-if="article.description" class="ng-binding ng-scope">...</p>
    <!-- end ngIf: article.description -->
    <span class="search-date ng-binding">May 8, 2018</span>
    <!-- ngIf: article.taxonomy -->
  </div>
  <!-- ngIf: article.image -->
```

PATRYCJA: setting up an app instance

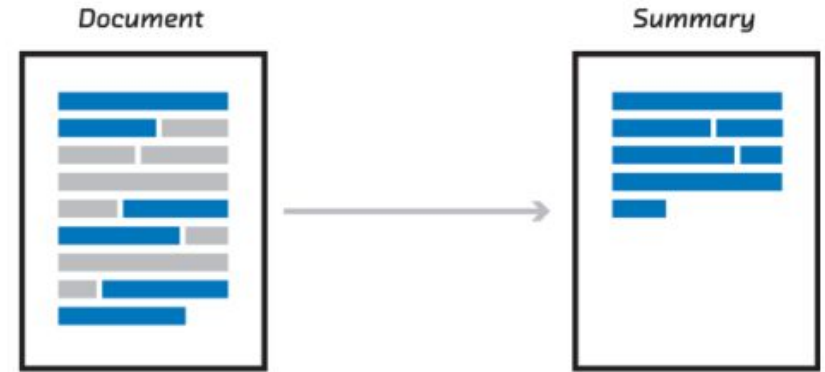
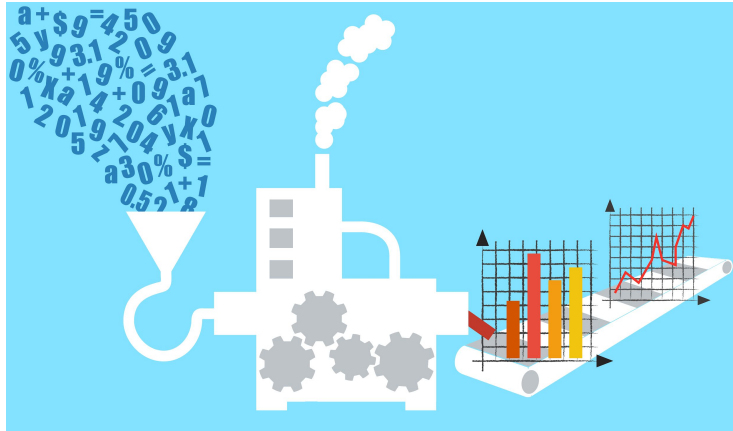


RICKY:

Interactive map of US in D3.js

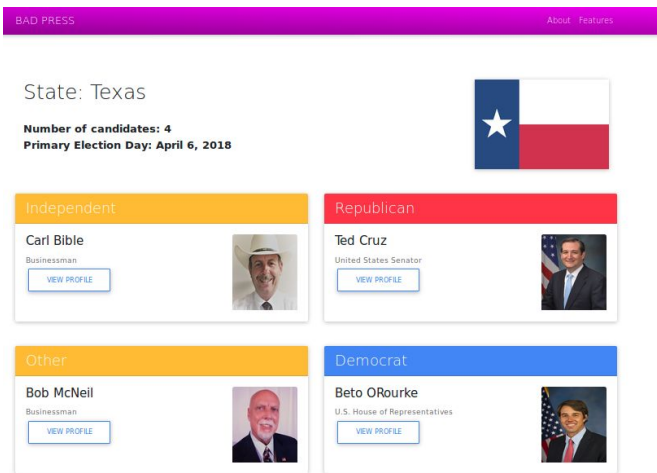


SARAH: cleaning data + summarizer

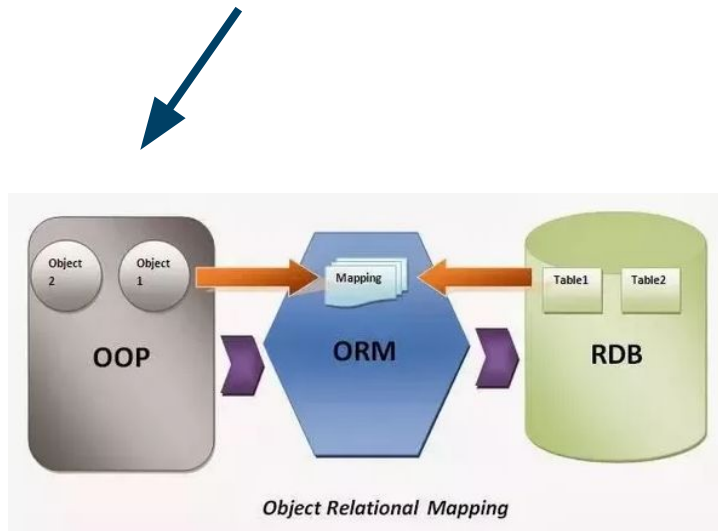


DJANGO CONTINUES..

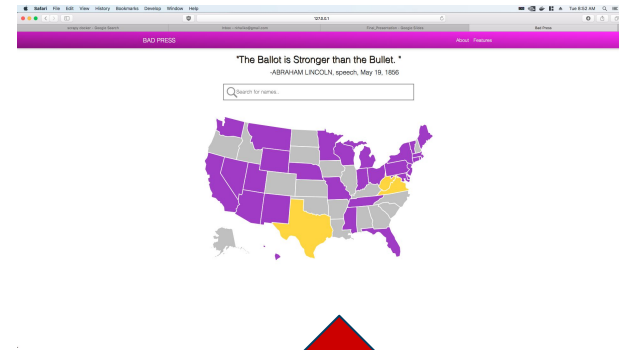
PATRYCJA :models and templating



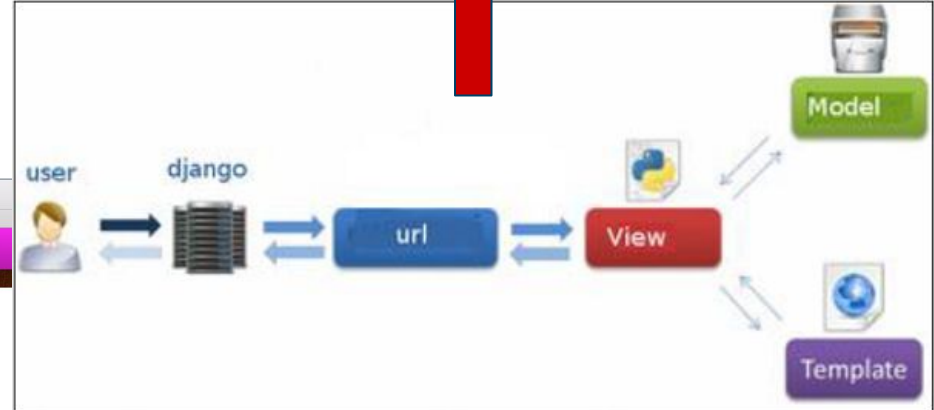
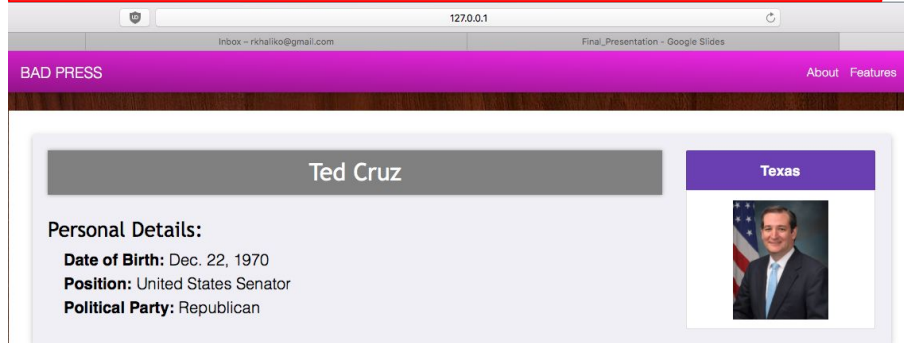
```
class Candidate(models.Model):
    """
    Model representing a candidate.
    """
    id = models.AutoField(primary_key=True)
    candidate_id = models.IntegerField(default=1)
    state = models.ForeignKey('State', on_delete=models.SET_NULL, null=True)
    name = models.CharField(max_length=200, null=True)
    first_name = models.CharField(max_length=100)
    last_name = models.CharField(max_length=100)
    date_of_birth = models.DateField(null=True, blank=True)
    place_of_birth = models.CharField(max_length=100)
    position = models.CharField(max_length=100)
    CANDIDATE_POSITION = (
        ('r', 'Republican'),
        ('d', 'Democrat'),
        ('i', 'Independent'),
        ('o', 'Other'),
    )
    party = models.CharField(max_length=1, choices=CANDIDATE_POSITION, blank=True)
    URL_photo = models.CharField(max_length=250)
    score_issue_1 = models.IntegerField()
    score_issue_2 = models.IntegerField()
    score_issue_3 = models.IntegerField()
    score_issue_4 = models.IntegerField()
    score_issue_5 = models.IntegerField()
```



RENAT:



<http://127.0.0.1:8000/badpress/candidate/Cruz/>

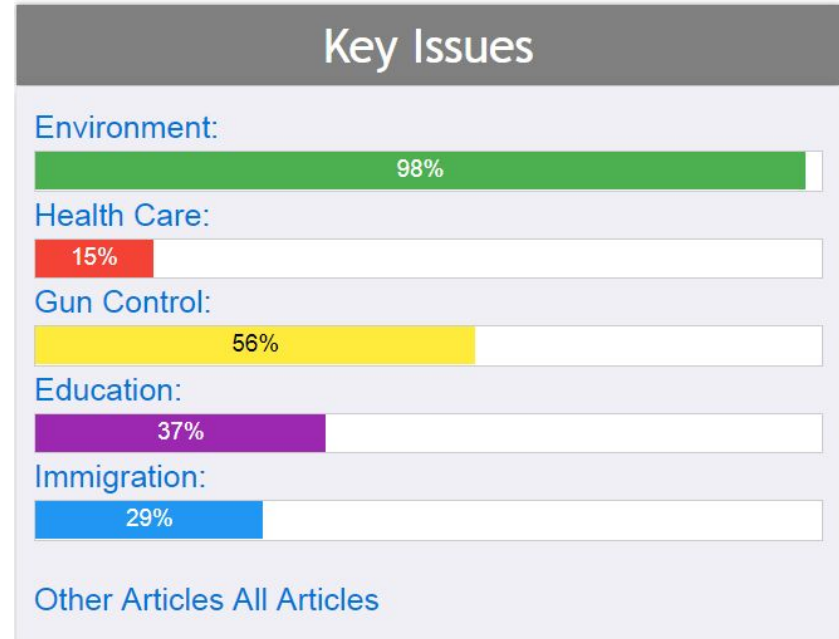


RICKY:

I used several web technologies, like Bootstrap, jquery, and d3. I also gained experience in the css libraries of Bootstrap and W3.

This experience taught me that sometimes, simpler is better.

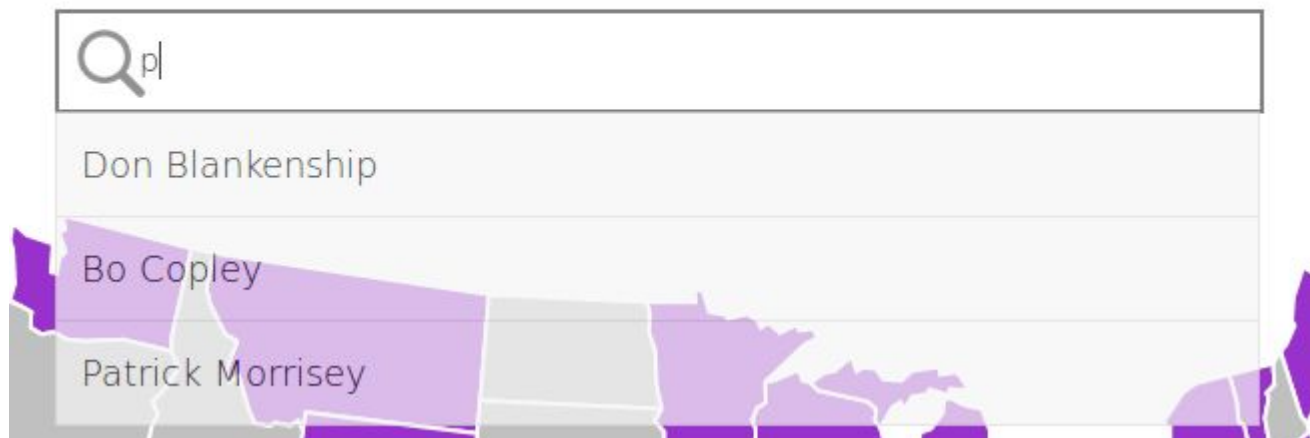
Key issues was made with d3 at first, but progress bars (css) achieved the same thing, with more style.



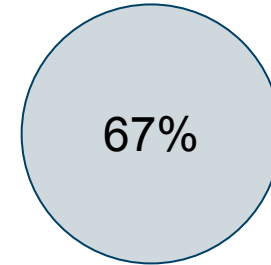
RICKY:

"The Ballot is Stronger than the Bullet. "

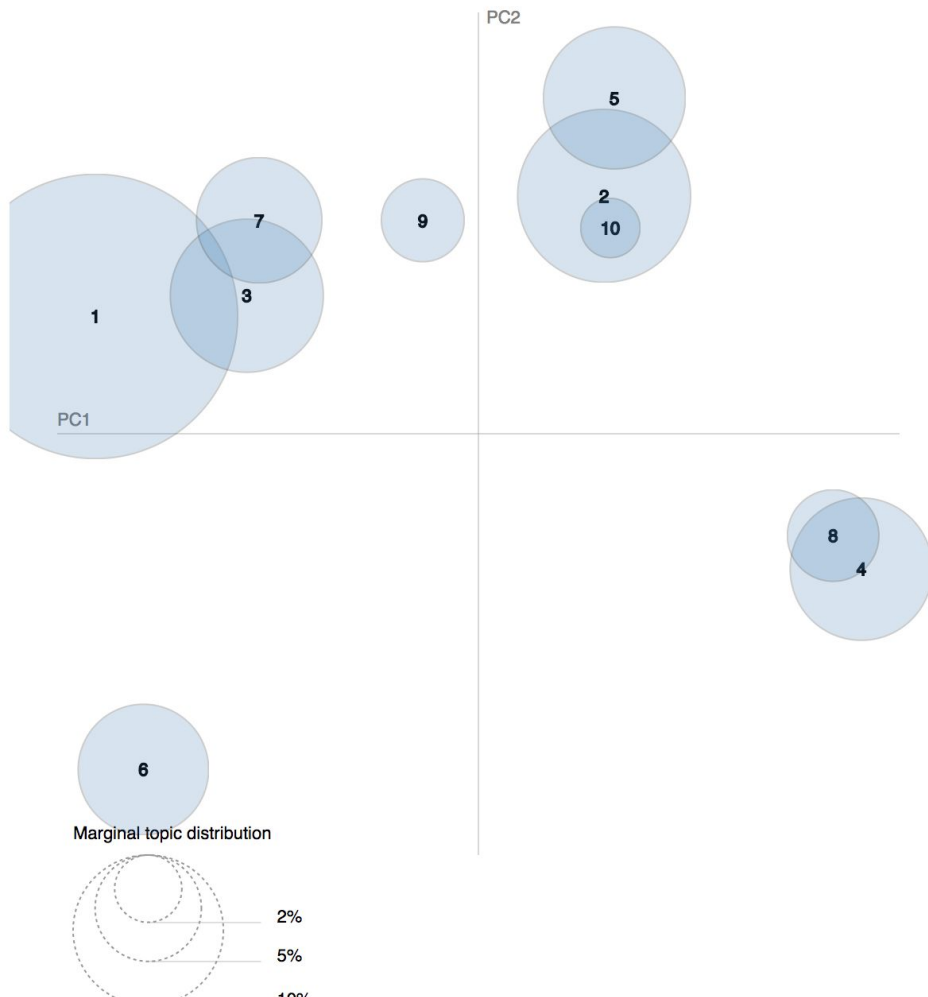
-ABRAHAM LINCOLN, speech, May 19, 1856



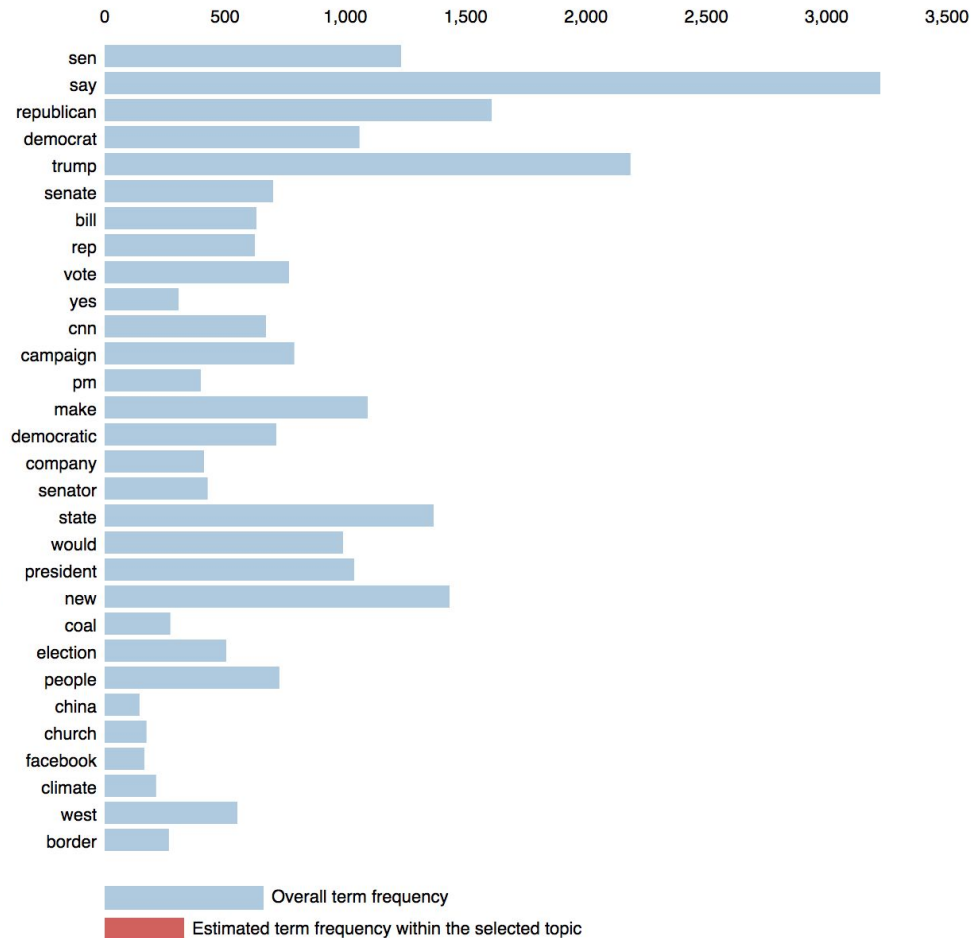
SARAH:



Intertopic Distance Map (via multidimensional scaling)



Top-30 Most Salient Terms⁽¹⁾



1. saliency(term w) = frequency(w) * [sum_t p(t | w) * log(p(t | w)/p(t))] for topics t; see Chuang et. al (2012)

2. relevance(term w | topic t) = λ * p(w | t) + (1 - λ) * p(w | t)/p(w); see Sievert & Shirley (2014)

LDA Analysis: First Pass, Top 30 words

july, june, say, aug, trump, sen, rep, festival

go, theater, street, campaign, art, border, republican

get, people, race, 212, museum, state, mr, democratic

music, pm, democrats, team, new, -, primary

ML

| | no stop words | stop words = basic | stop words = extra | stop words = basic+extra |
|------------------------|---------------|--------------------|--------------------|--------------------------|
| Multinomial NaiveBayes | 65.3% | 66.7% | 65.3% | 65.3% |
| SVM | 78.9% | 78.2% | 78.2% | 77.6% |

ML

stemming

| | no stop words | stop words = basic | stop words = extra | stop words = basic+extra |
|------------------------|---------------|--------------------|--------------------|--------------------------|
| Multinomial NaiveBayes | 68.0% | 70.06% | 68.0% | 69.4% |
| SVM | 81.0% | 81.0% | 81.0% | 81.0% |

ML

lemmatize

| | no stop words | stop words = basic | stop words = extra | stop words = basic+extra |
|------------------------|---------------|--------------------|--------------------|--------------------------|
| Multinomial NaiveBayes | 67.3% | 68.7% | 67.3% | 68.7% |
| SVM | 78.9% | 81.0% | 82.3% | 82.3% |

ML

stemming + lemmatize

| | no stop words | stop words = basic | stop words = extra | stop words = basic+extra |
|------------------------|---------------|--------------------|--------------------|--------------------------|
| Multinomial NaiveBayes | 68.7% | 69.4% | 68.0% | 68.0% |
| SVM | 80.3% | 80.1% | 79.6% | 79.6% |

WHAT WENT WELL?

- Web scraping 3 news sources
- Connecting data from database to website using Django
- Templating the pages
- Making data pieces on candidate profile

WHAT DID NOT GO WELL?

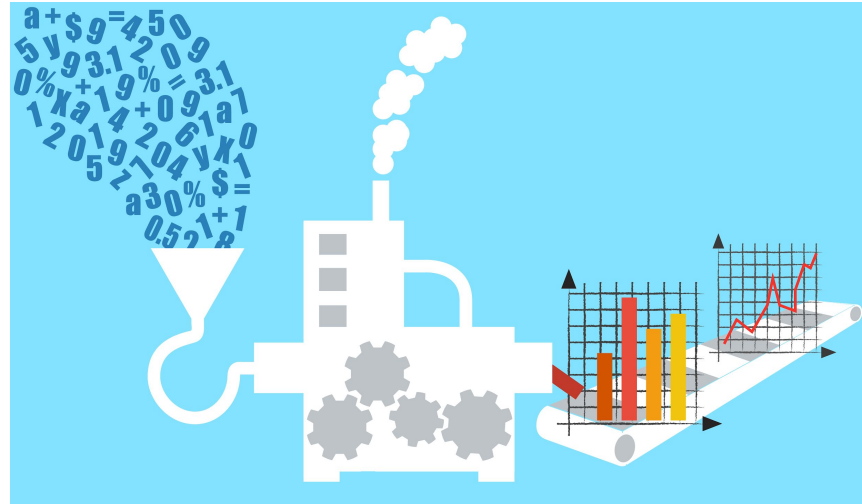
- Scrapy-Splash optimization issues on Fox News Website
- Change Sqlite to MySQL
- Lack of articles about some of the candidates (Assumptions)
- Polish of website--not standardized for all devices

WHAT DID WE LEARN?

- How websites work
- Team work
- GitHub
- How to use ORM

WHAT DID WE LEARN?

SARAH:



Sentiment Analysis in the News

**Alexandra Balahur¹, Ralf Steinberger², Mijail Kabadjov², Vanni Zavarella²,
Erik van der Goot², Matina Halkia², Bruno Pouliquen², Jenya Belyaeva²**

¹ University of Alicante, Department of Software and Computing Systems

Ap. de Correos 99, E-03080 Alicante, Spain

² European Commission – Joint Research Centre

IPSC - GlobeSec - OPTIMA (OpenSource Text Information Mining and Analysis)

T.P. 267, Via Fermi 2749

21027 Ispra (VA), Italy

abalahur@dlsi.ua.es,

{Ralf.Steinberger, Mijail.Kabadjov, Erik.van-der-Goot, Matina.Halkia, Bruno.Pouliquen}@jrc.ec.europa.eu,
{Vanni.Zavarella, Jenya.Belyaeva}@ext.jrc.ec.europa.eu

Abstract

Recent years have brought a significant growth in the volume of research in sentiment analysis, mostly on highly subjective text types (movie or product reviews). The main difference these texts have with news articles is that their target is clearly defined and unique across the text. Following different annotation efforts and the analysis of the issues encountered, we realised that news opinion mining is different from that of other text types. We identified three subtasks that need to be addressed: definition of the target; separation of the good and bad news content from the good and bad sentiment expressed on the target; and analysis of clearly marked opinion that is expressed explicitly, not needing interpretation or the use of world knowledge. Furthermore, we distinguish three different possible views on newspaper articles – author, reader and text, which have to be addressed differently at the time of analysing sentiment. Given these definitions, we present work on mining opinions about entities in English language news, in which (a) we test the relative suitability of various sentiment dictionaries and (b) we attempt to separate positive or negative opinion from good or bad news. In the experiments described here, we tested whether or not subject domain-defining vocabulary should be ignored. Results showed that this idea is more appropriate in the context of news opinion mining and that the approaches taking this into consideration produce a better performance.

Large-Scale Sentiment Analysis for News and Blogs

Namrata Godbole*

namratagodbole@gmail.com

Manjunath Srinivasaiah*

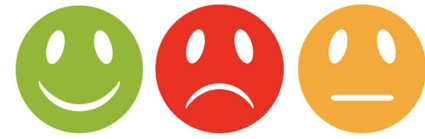
manj.blr@gmail.com

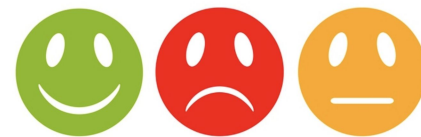
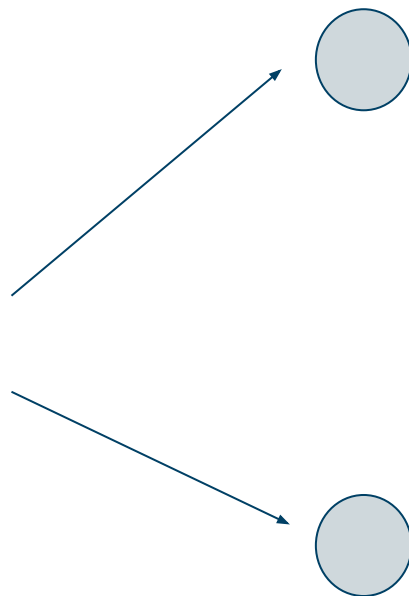
Steven Skiena[◇]

skiena@cs.sunysb.edu

*Google Inc., New York NY, USA

[◇]Dept. of Computer Science, Stony Brook University, Stony Brook, NY 11794-4400, USA





WHAT DID WE LEARN?

SARAH:

— Plans change

Environment
Healthcare
Gun Control
Education
Immigration
Other



Environment
Healthcare
Gun Control
Election
Immigration
Other