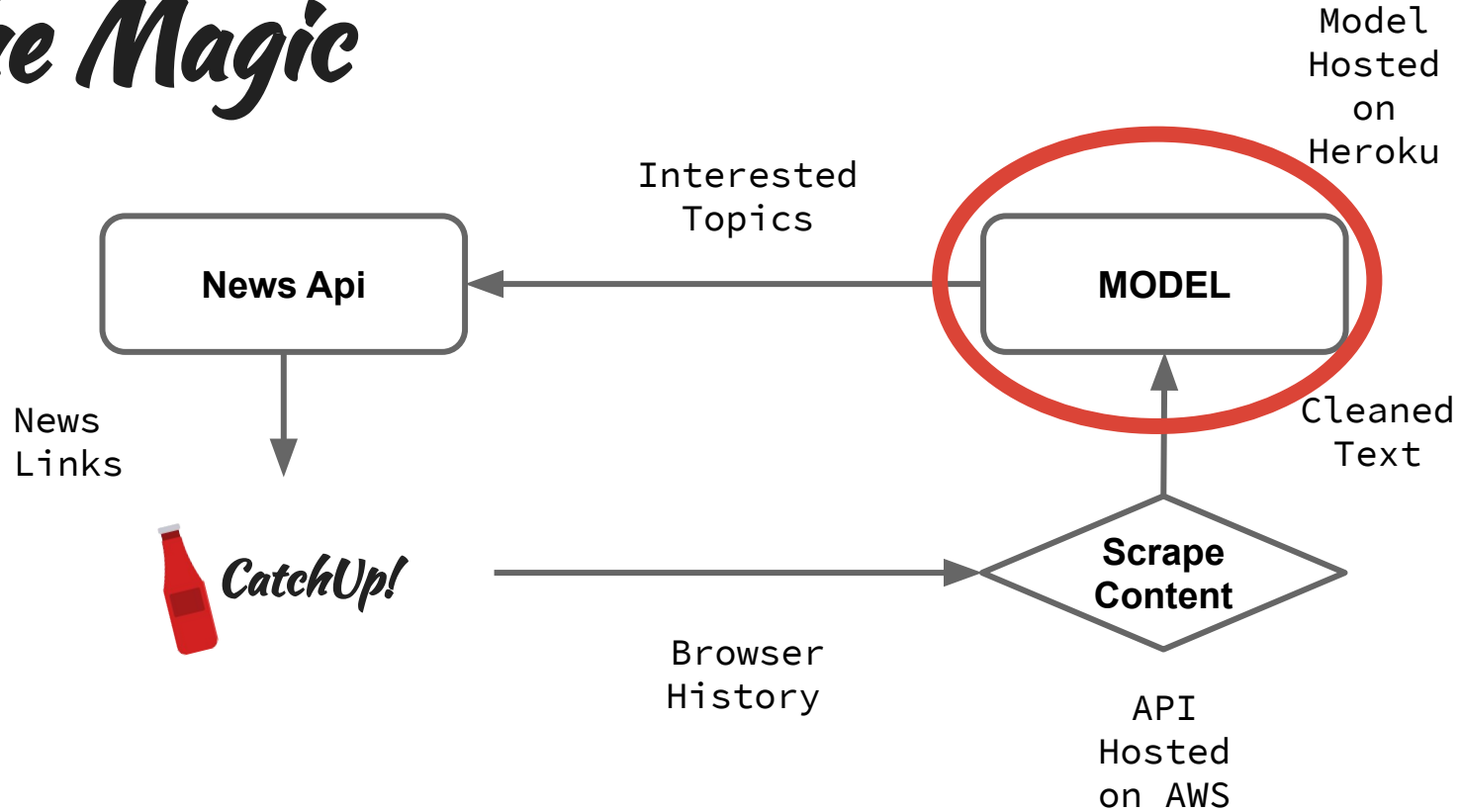




News Recommendation Chrome Extension

We need our news
FRESH
RELEVANT
EXCITING
EASY

The Magic



Datasets

Harvard University (3824 Articles)

24	2017/2/7	http://www.cnn.com/2017/02/07/politics/iran-khamenei-trump-comments/index.html	Trump 'shows true face of US,' says Iran's Supreme Leader	(CNN)Iran's Supreme Leader Ayatollah Ali Khamenei said Tuesday that he's grateful to President Donald Trump for showing "the true face of the US."
25	2017/2/7	http://www.cnn.com/2017/02/07/politics/yemen-raid-al-qaeda/index.html	Al Qaeda leader mocks Trump after Yemen raid	(CNN)Days after a raid on an al Qaeda compound in Yemen led to the first US military combat death under Donald Trump, the leader of al Qaeda in the Arabian Peninsula released an audio message taunting the new American President.
26	2017/2/7	http://www.cnn.com/2017/02/07/politics/trump-japan-shinzo-abe-mar-a-lago/index.html	Trump teeing up a softer diplomatic approach with Japan	Washington (CNN)President Donald Trump is turning on the charm. Or at least, he's turning on his golf cart.

Datasets

Webhose (170882 Articles)

```
[{'organizations': [],  
  
  'uuid': 'b166f6b85b58506e290687e3b4e1273b57004816',  
  
  'thread': {'social': {'gplus': {'shares': 0},  
  
    'pinterest': {'shares': 0},  
  
    'vk': {'shares': 0},  
  
    'linkedin': {'shares': 0},  
  
    'facebook': {'likes': 287, 'shares': 287, 'comments': 0},  
  
    'stumbledupon': {'shares': 0}}},
```

Datasets

Merged the datasets

content

0 <http://abcnews.go.com/Politics/pence-break-tie...>

1 <http://abcnews.go.com/Politics/wireStory/melan...>

2 <http://abcnews.go.com/Politics/wireStory/trump...>

3 <http://abcnews.go.com/Politics/appeals-court-d...>

4 <http://abcnews.go.com/US/23-states-winter-weat...>

...

...

Data Cleaning

<http://www.cnn.com/2017/02/07/politics/barack-obama-kitesurfing-richard-branson/index.html> Barack Obama's kitesurfing adventure with Richard Branson
Before returning to Washington, the former president dove into an aquatic and athletic challenge with his friend, ere ..

Change to
lowercase.

Remove all
characters

http www cnn com politics barack obama
kitesurfing richard branson index html barack
obamas kitesurfing adventure with richard
branson
before returning to washington the former
president dove into an aquatic and athletic
challenge with his friend ere ..

Remove
stopwords
Lemmatizing

Text
Vectorize

[(2,1),(3,2),(4,2)]

[politics barack obama kitesurfing richard branson
barack obamas kitesurfing adventure richard
branson]

Latent Dirichlet Allocation

LDA

Number of Topics	10	15	20
Coherence Score	0.45	0.48	0.47

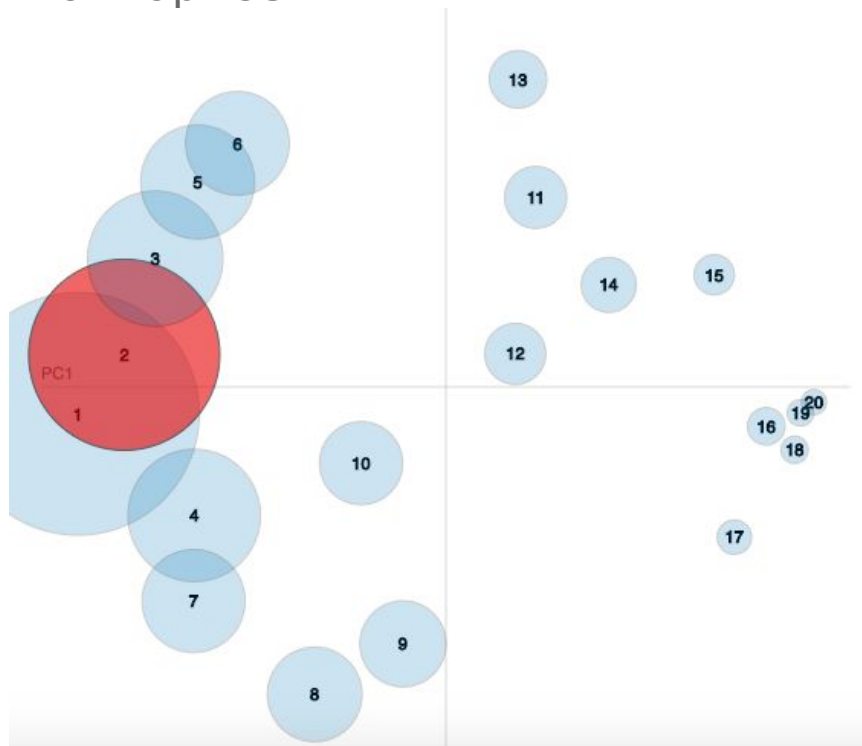
Coherence Score tells you how different are the topics clusters.
The higher the score means the topics are better separated and thus better.

Football and Soccer. Bad score.

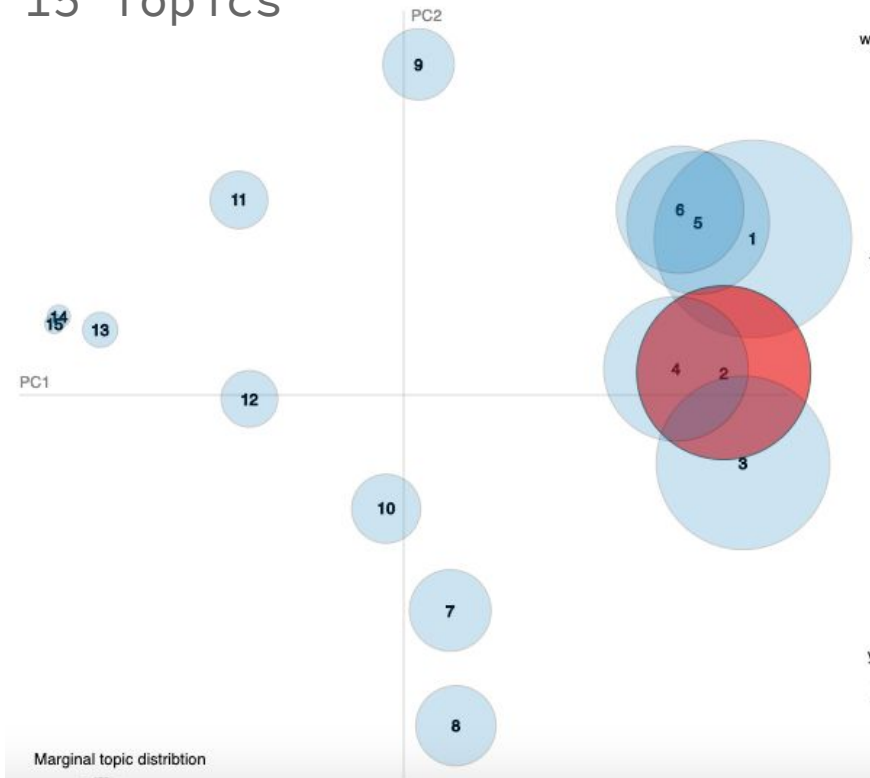
Swimming and Finance. Good score.

Visualisation

20 Topics



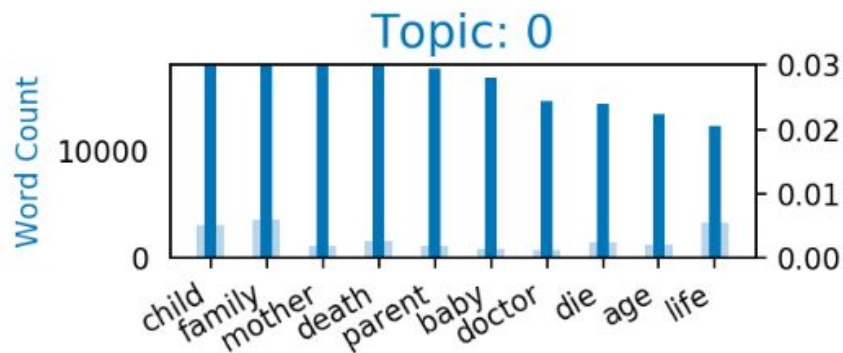
15 Topics



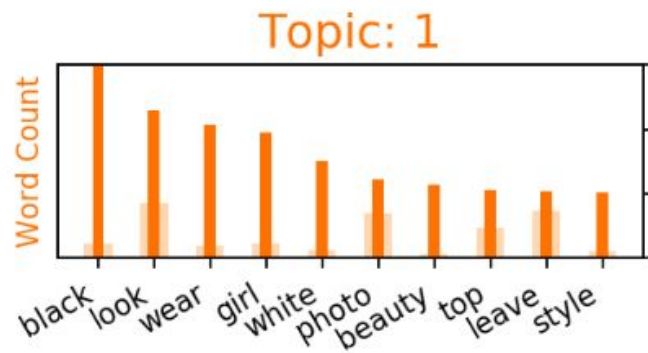
Article's Dominant Topic

Document_No	Dominant_Topic	Topic_Perc_Contrib	Keywords	Text	
0	0	18.0	0.2733	phone, feature, user, online, launch, device, ...	http://www.phonearena.com/news/LG-V30-rumored-...
1	1	0.0	0.3309	image, music, year, world, art, space, perform...	http://pitchfork.com/reviews/albums/22948-the-...
2	2	8.0	0.2916	country, government, force, military, official...	https://www.rt.com/news/381216-trump-merkel-na...
3	3	17.0	0.2444	family, child, life, home, live, friend, die, ...	http://www.self.com/story/mark-zuckerberg-pris...
4	4	4.0	0.1371	company, work, business, job, project, plan, g...	http://www.news.com.au/finance/business/retail...
5	5	10.0	0.1217	trump, vote, election, president, campaign, st...	http://www.smh.com.au/federal-politics/politic...
6	6	7.0	0.4666	pay, money, government, cost, percent, plan, c...	http://www.cbc.ca/news/canada/saskatchewan/sas...
7	7	1.0	0.4077	photo, show, share, wear, reveal, picture, lov...	http://www.mirror.co.uk/3am/celebrity-news/jer...
8	8	12.0	0.2902	people, school, student, country, group, commu...	http://www.nbcnews.com/news/latino/opinion-ice...
9	9	3.0	0.2257	film, show, good, story, movie, award, watch, ...	http://www.hellomagazine.com/celebrities/20170...

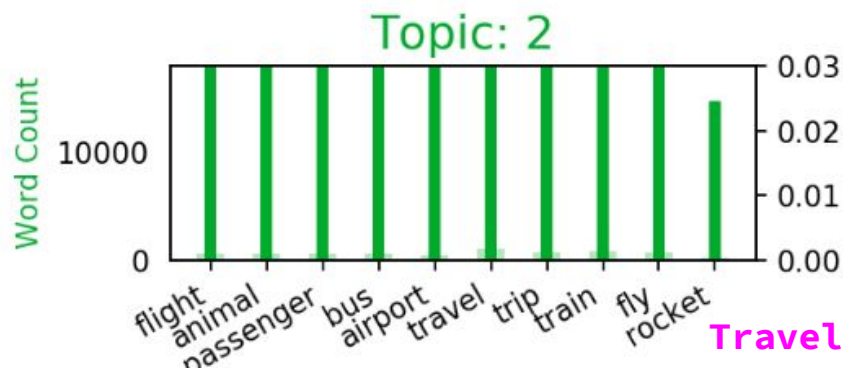
Topic Inference



Medical



Fashion



Travel



Sports

Classification Model

Linear Regression	Naive Bayes
0.745	0.740

The Magic

