



# Detecting the ‘Fake News’ Before It Was Even Written

Preslav Nakov

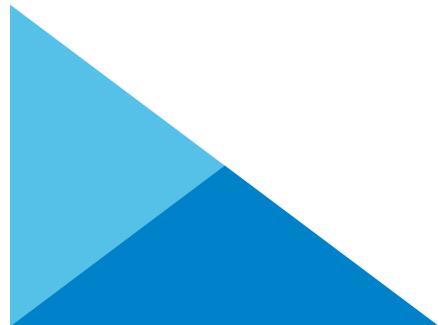
Qatar Computing Research Institute, HBKU

*Singapore Symposium on Natural Language Processing (SSNLP 2019)*

*October 31, 2019*

*Singapore*

# What is “Fake News”?



08.03.2019 BNT News

## Mariya Gabriel: 75% of young people do not recognize fake news

Europe is taking small steps to fight against fake news, but the road is very long. Mariya Gabriel, European Commissioner for Digital Economy and Digital Society told ...

Share with friends



<https://www.bnt.bg/en/a/mariya-gabriel-75-of-young-people-do-not-recognize-fake-news>

**Monday, December 12, 2016**

## **WAR ON CHRISTMAS: OBAMA ORDERS BAN ON CHRISTMAS CARDS TO MILITARY**

The image is a composite of two photographs. On the left, a female news anchor with blonde hair, wearing a dark sleeveless top, is speaking on a television news set. On the right, a photograph of a newspaper front page from 'DAILY EXPRESS' is shown. The headline on the paper reads 'CHRISTMAS IS BANNED: IT OFFENDS MUSLIMS'. Below the headline, there is a sub-headline 'OBAMA ORDERS BAN ON ALL CHRISTMAS CARDS TO OVERSEAS SERVICEMEN' and a smaller text 'PENTAGON TO CENSOR ALL OUTGOING MAIL THAT APPEARS TO BE CHRISTMAS THEMED'. The background of the entire composite image is a blue and purple gradient.

**OBAMA ORDERS BAN ON ALL CHRISTMAS CARDS TO OVERSEAS SERVICEMEN**

**PENTAGON TO CENSOR ALL OUTGOING MAIL THAT APPEARS TO BE CHRISTMAS THEMED**

Thursday 11 April 2019 by Chris Ballard

## TV fans delighted as Brexit renewed for another season



Like Bozhidar Bozhanov and 358K others like this.

Fans of Britain's long-running comedy-drama Brexit are today overjoyed after learning that another season of the popular show has been commissioned.

Just when it looked like the UK's telly addicts would need to find another boxset to binge it has been announced that Brexit – possibly the greatest farce since *Fawlty Towers* and a fantasy to rival *Game of Thrones* – will roll on for at least one more season.



Brexit aficionado Simon Williams had been bracing himself for the show's emotional finale.

"I'm so happy!" he said. "There's still so much potential with these characters, so many more votes they could have on Theresa May's deal."

≡ SECTIONS

WATCHLIST >

MarketWatch

# Web creator Tim Berners-Lee blasts Facebook, saying it makes his invention easy to 'weaponize'

Published: Mar 19, 2018 4:40 a.m. ET



Aa Aa

*On World Wide Web's 29th birthday, Tim Berners-Lee criticizes its "gatekeepers" such as Facebook and suggests more regulation*



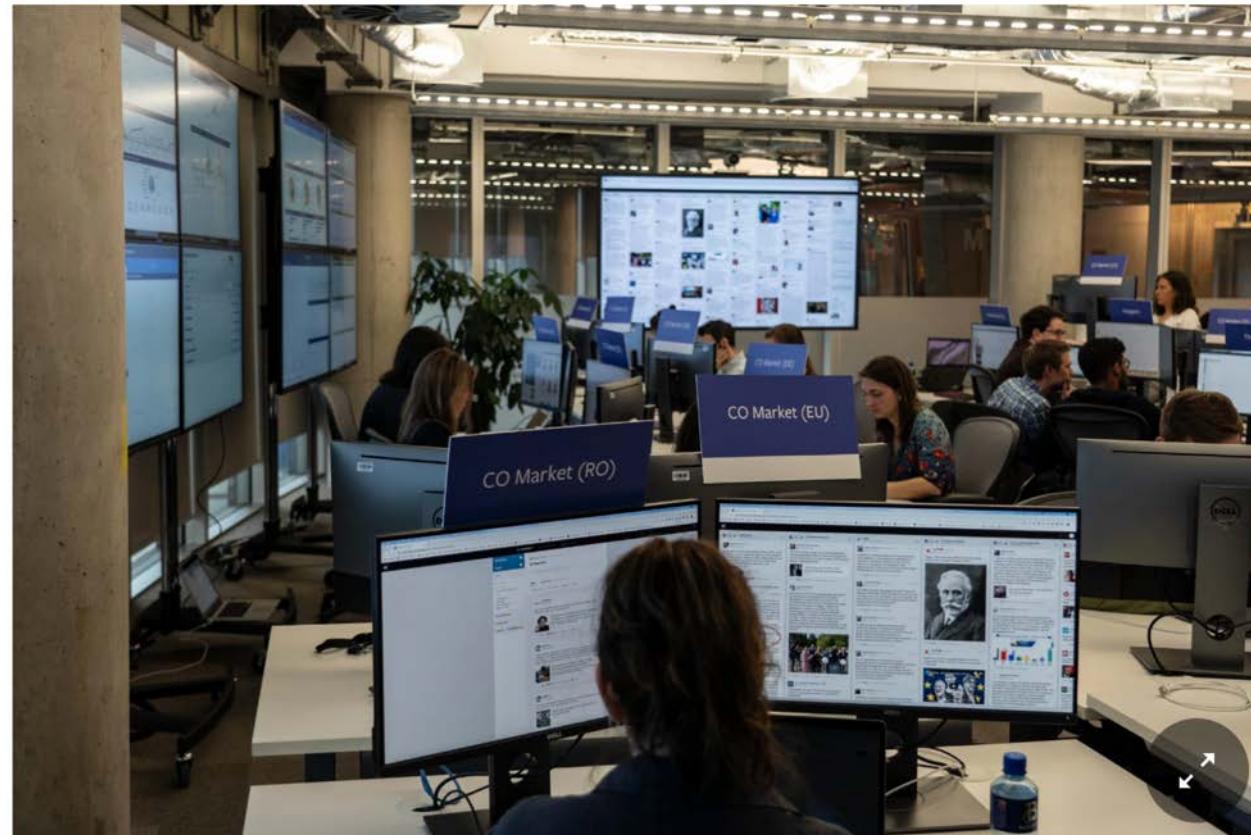
Getty Images

Computer scientist Tim Berners-Lee, the "father of the World Wide Web."

<https://www.marketwatch.com/story/webs-creator-blasts-the-tech-giants-that-make-his-invention-easy-to-weaponize-2018-03-12>

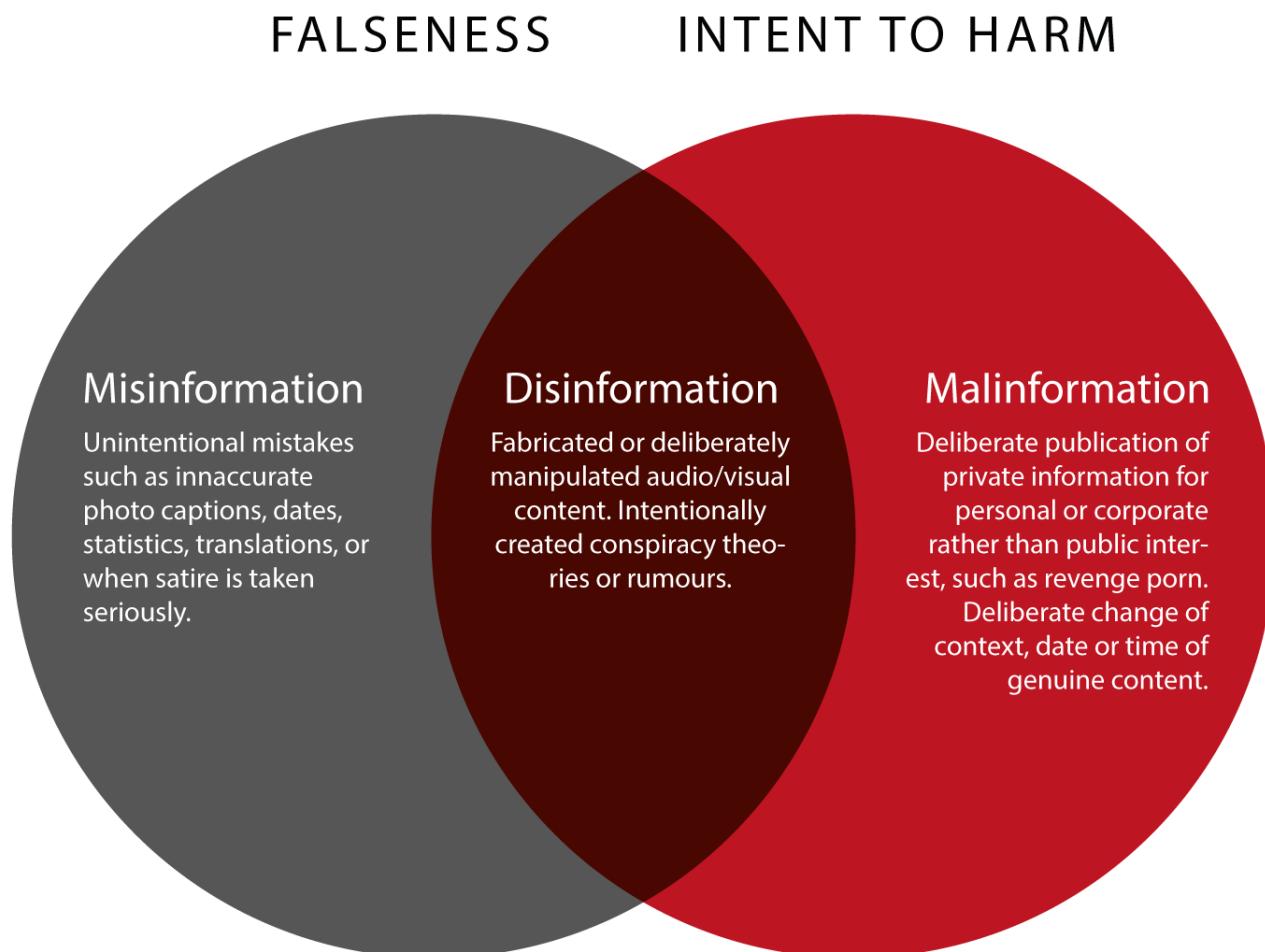
The New York Times

# *At Least 70 Countries Have Had Disinformation Campaigns, Study Finds*



<https://www.nytimes.com/2019/09/26/technology/government-disinformation-cyber-troops.html>

# TYPES OF INFORMATION DISORDER



# Is Fact-Checking the Solution to Disinformation?

# A Lot of Work Focused on Fact-Checking *Claims, Rumors, Articles*



Elon Musk

@elonmusk

Following

Am considering taking Tesla private at \$420. Funding secured.

9:48 AM - 7 Aug 2018

## 'Putin is DEAD' Shock claims Vlad was killed YEARS ago and Russia led by 'body double'

Vladimir Putin was killed in a plot by the CIA and MI6 and replaced with a body double who still leads Russia, according to astonishing claims.

### During Farewell Address to Armed Forces, Obama URGES Soldiers to Do the Unthinkable to Trump

Like

Share

945 people like this. Be the first of your friends.



Do you agree with the Senate's vote on Kavanaugh?

Yes

No

Not sure

2,764 Votes • Promoted Content

There is no question this country has been infiltrated by the far-left. If you were expecting President Obama to engage in a civil discourse with the president-elect, then I am afraid you are sadly mistaken.

# But Can We Fact-Check Every Claim in the World?



# Fake News Generation Can Now Be Automated

GPT-2: <https://taktotransformer.com>

Custom prompt

The price of oil just hit \$100 per barrel.

GENERATE ANOTHER

## Completion

**The price of oil just hit \$100 per barrel.**

"That's a price that we had not anticipated."

"I think we've got a pretty good idea as to what the market situation is going on," said Mr. Treloar.

"We can tell that the market is still holding at the high end of its range and we think it could continue to drop over the course of the remainder of the year or maybe for the rest of the year."

"We certainly look for oil to be in the \$60s to be around now."

Grover: <https://grover.allenai.org>

Oil is back above \$100. That should shock you into thinking more about energy policy. Or it should, but in fact it doesn't.

Yes, we might turn off the heat if we had to – but there's no way we can walk away from oil. It's an essential part of everyday life in almost all parts of the world.

And so we will continue to pump it. Probably, the oil-supply cartel OPEC (which always has good sense – remember this?).

That's going to be bad news for consumers everywhere, but especially in the energy-importing countries where more than half of our oil imports come from. Some of these markets are in major trouble, including Poland, India, and the East European region, but many other countries – led by Russia, Venezuela,

Iran, and Nigeria – are getting richer and will oil remains affordable.

Take a quiz:

<https://quiz.newsyoucantuse.com/>

THE DAILY NEWSLETTER  
Sign up to our daily email newsletter

# NewScientist

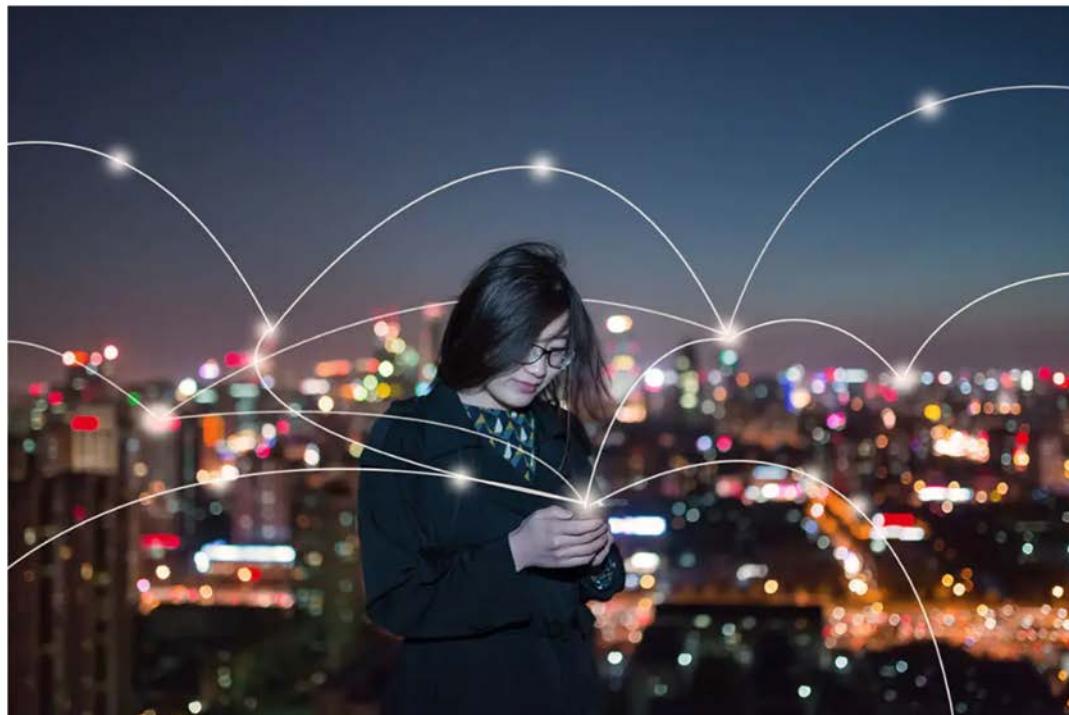
News **Technology** Space Physics Health Environment Mind Video | Tours Events Jobs

## Fake news travels six times faster than the truth on Twitter



TECHNOLOGY 8 March 2018

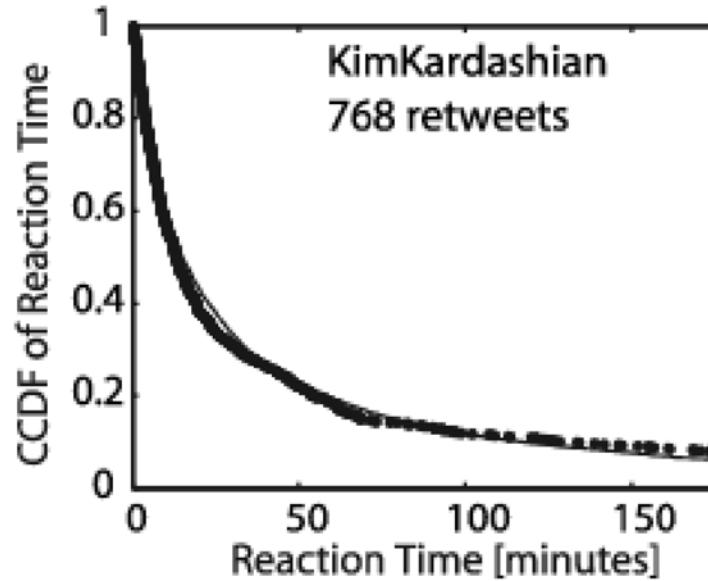
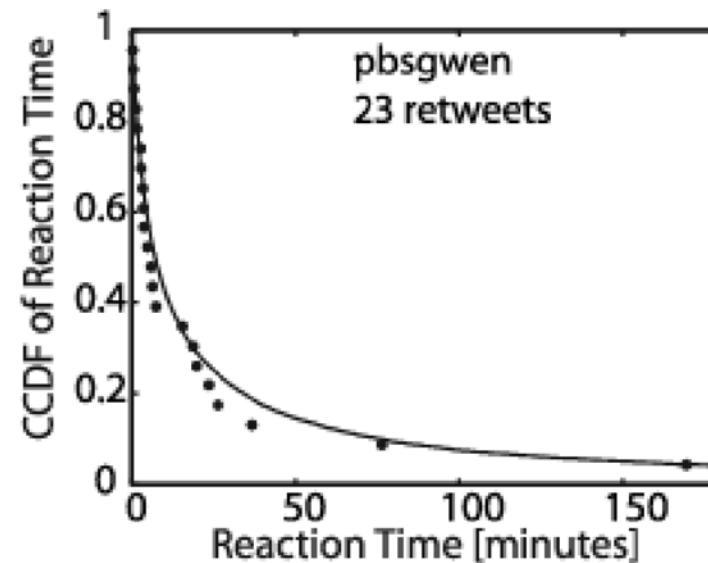
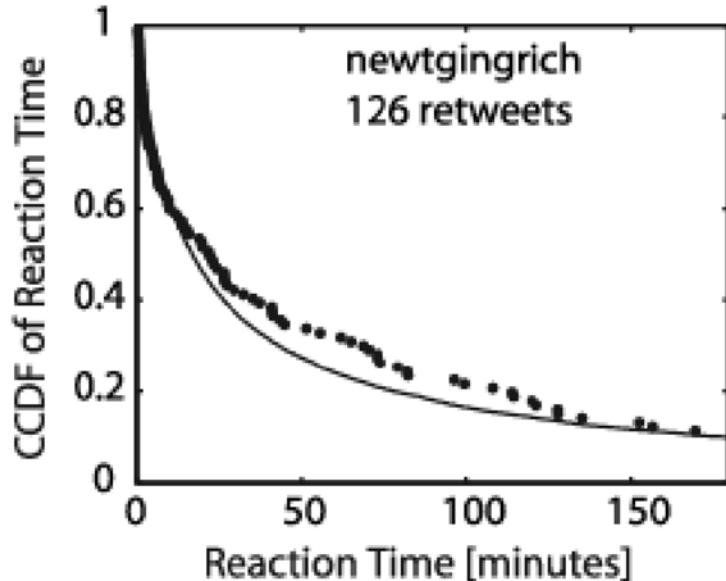
By [Chris Stokel-Walker](#)



News travels fast, especially when it's wrong

<https://www.newscientist.com/article/2163226-fake-news-travels-six-times-faster-than-the-truth-on-twitter/>

**50% of the spread  
of "fake news"  
on Twitter:  
<10 minutes**



Tauhid Zaman, Emily B. Fox, and Eric T. Bradlow  
A Bayesian approach for predicting the popularity of  
tweets.  
Ann. Appl. Stat. 8(3):1583-1611, 2014.  
<https://projecteuclid.org/euclid.aoas/1414091226>

# So, Better to Go After the Source

<http://www.angrypatriotmovement.com/>



Search the site



Would you like to retire rich?

Yes

During Farewell Address to Armed Forces, Obama URGES Soldiers to Do the Unthinkable to Trump



Like



Share

945 people like this. Be the first of your friends.



Do you agree with the Senate's vote on Kavanaugh?

Yes

No

Not sure

2,764 Votes · Promoted Content

There is no question this country has been infiltrated by the far-left. If you were expecting President Obama to engage in a civil discourse with the president-elect, then I am afraid you are sadly mistaken.

**Thus, We Can Detect “Fake News”  
Before It Was even Written!**



# Can We Win the War on “Fake News”?

# Complex Problem, No Easy Solution

- Need for cooperation
  - social media, technology companies
  - governments
  - international organizations
  - civil society: journalists, fact-checkers, media, NGOs, etc.
  - researchers: academia, industry



# Finland is winning the war on fake news. What it's learned may be crucial to Western democracy

By Eliza Mackintosh, CNN

Video by Edward Kiernan, CNN



<https://edition.cnn.com/interactive/2019/05/europe/finland-fake-news-intl/>



*“Propaganda becomes ineffective the moment we are aware of it.”*

Joseph Goebbels (1897-1945)

# The *Tanbih* Mega-Project

# *Tanbih*: Raising Awareness



## Highlights:

- **Disinformation-aware news aggregator**
- **Media profiles:** help fact-check the news before it was even written
- **Fine-grained propaganda analysis:** trains people to recognize it

## BUSINESS

SPORTS

## POLITICS

TECH & SCI



"We're OK" says British colleague of eight climbers feared dead in India

1

1 of 2

NODE



## 'Nothing Short of Madness': Netanyahu Accused of Trying to Turn Israel Into Iran

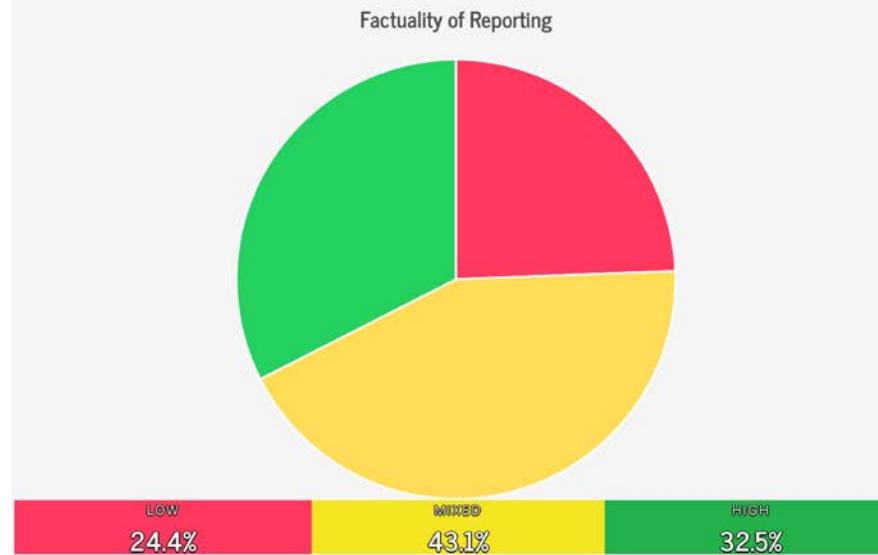
1

1 of 2

MIDDLEEAST

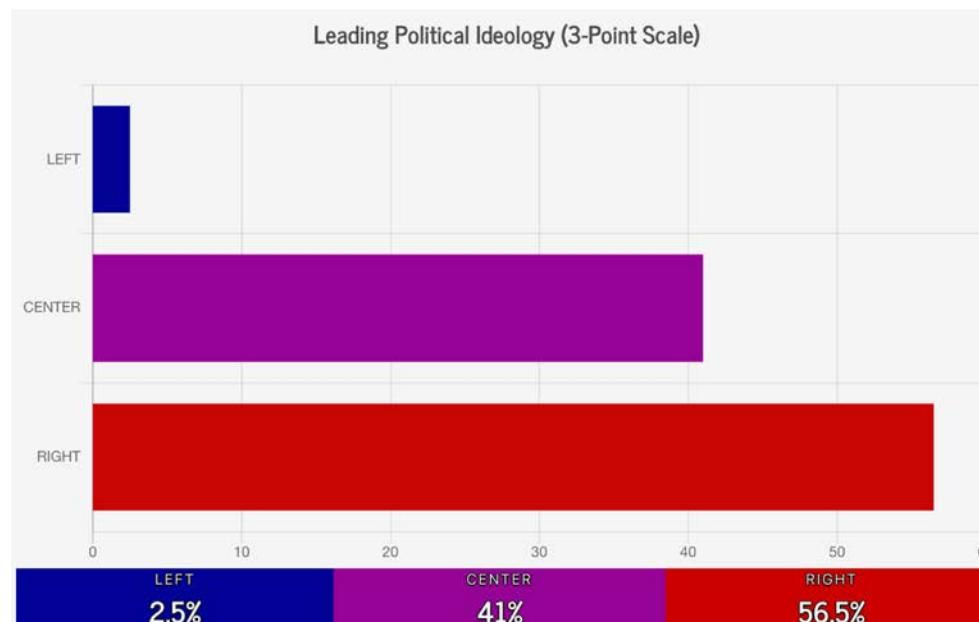


Try Tanbih:  
<http://www.tanbih.org>



We use AI to generate an automatic estimation of the factuality of reporting for the target news outlet. We show to what degree the model thinks that the target news outlet is likely to have low vs. mixed. vs. high factuality of reporting. This is estimated based on variety of information sources: a sample of articles published by the target news outlet, information from its Wikipedia page (if any), from its Twitter account (if any), from Web traffic, and from the structure of its URL.

[Learn more](#)



We use AI to generate an automatic estimation of the leading political ideology for the target news outlet. We show to what degree the model thinks that the target news outlet is likely to be left vs. centric vs. right leaning (on a 3-point scale). This is estimated based on variety of information sources: a sample of articles published by the target news outlet, information from its Wikipedia page (if any), from its Twitter account (if any), from Web traffic, and from the structure of its URL.

[Learn more](#)

# **Media Bias and Factuality of Reporting**

# MEDIA BIAS/FACT CHECK

The Most Comprehensive Media Bias Resource



SEARCH ONLY MEDIA SOURCES (best results)



[HOME](#)   [SEARCH](#)   [ABOUT](#)   [METHODOLOGY](#)   [MBFC NEWS](#)   [ORIGINAL ARTICLES/NEWS](#)   [LIVE TV NEWS](#)   [APPS/EXTENSIONS](#)  
[SUBMIT SOURCE](#)   [SUBMIT FACT CHECK](#)   [SOURCES PENDING](#)   [FACTUAL NEWS SEARCH](#)   [FILTERED SEARCH](#)   [RSS](#)

(NEW) [HELP US FACT CHECK](#)

[Left Bias](#)   [Left-Center Bias](#)   [Least Biased](#)   [Right-Center Bias](#)   [Right Bias](#)   [Pro-Science](#)   [Conspiracy-Pseudoscience](#)   [Questionable Sources](#)

[Satire](#)   (NEW) [Re-Evaluated Sources](#)

We are the most comprehensive media bias resource on the internet. There are currently 2500+ media sources listed in our database and growing every day. Don't be fooled by Fake News sources. Use the search feature above (Header) to check the bias of any source. Use name or url.

LATEST

Select Language ▾

Powered by [Google Translate](#)

[Donate](#)

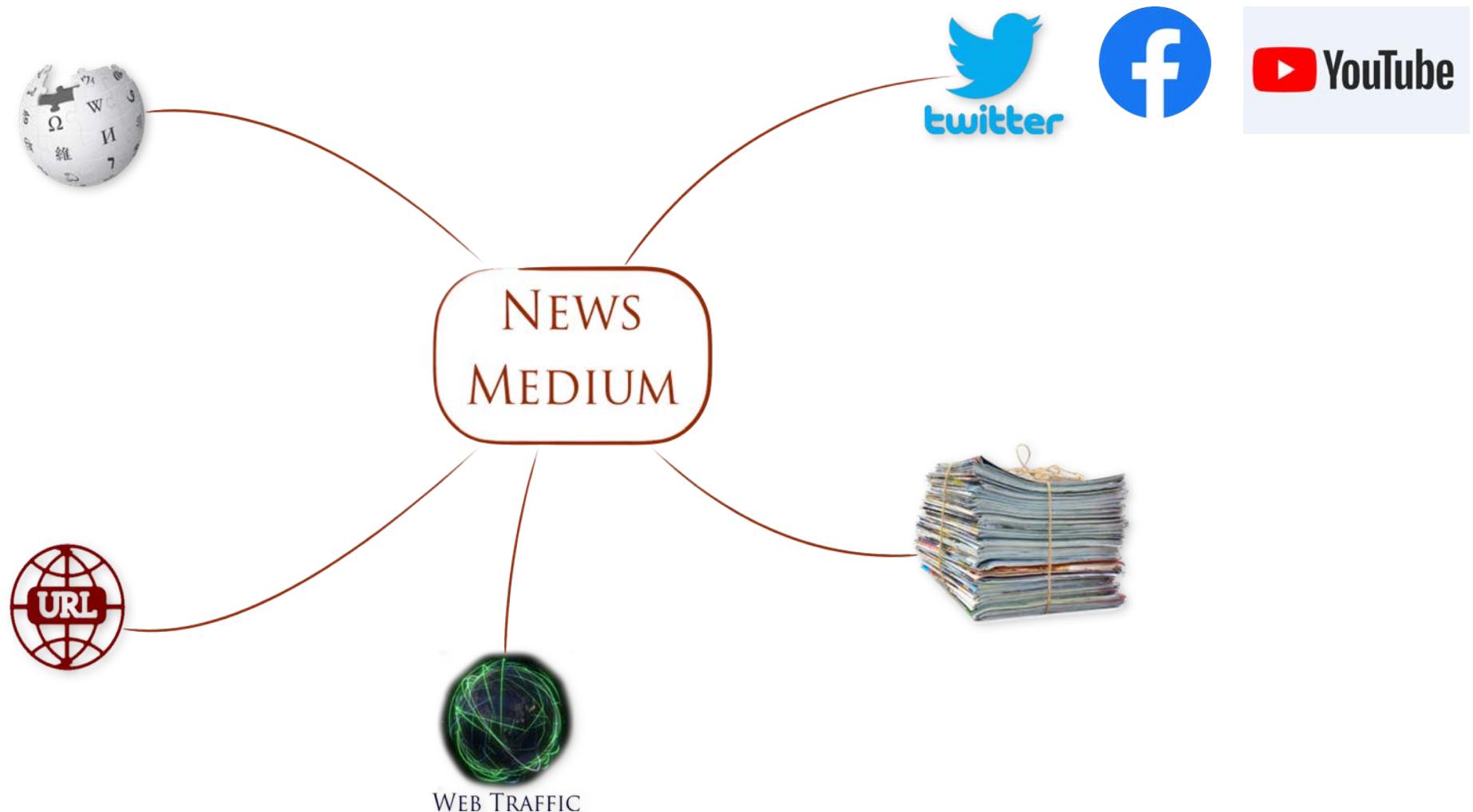


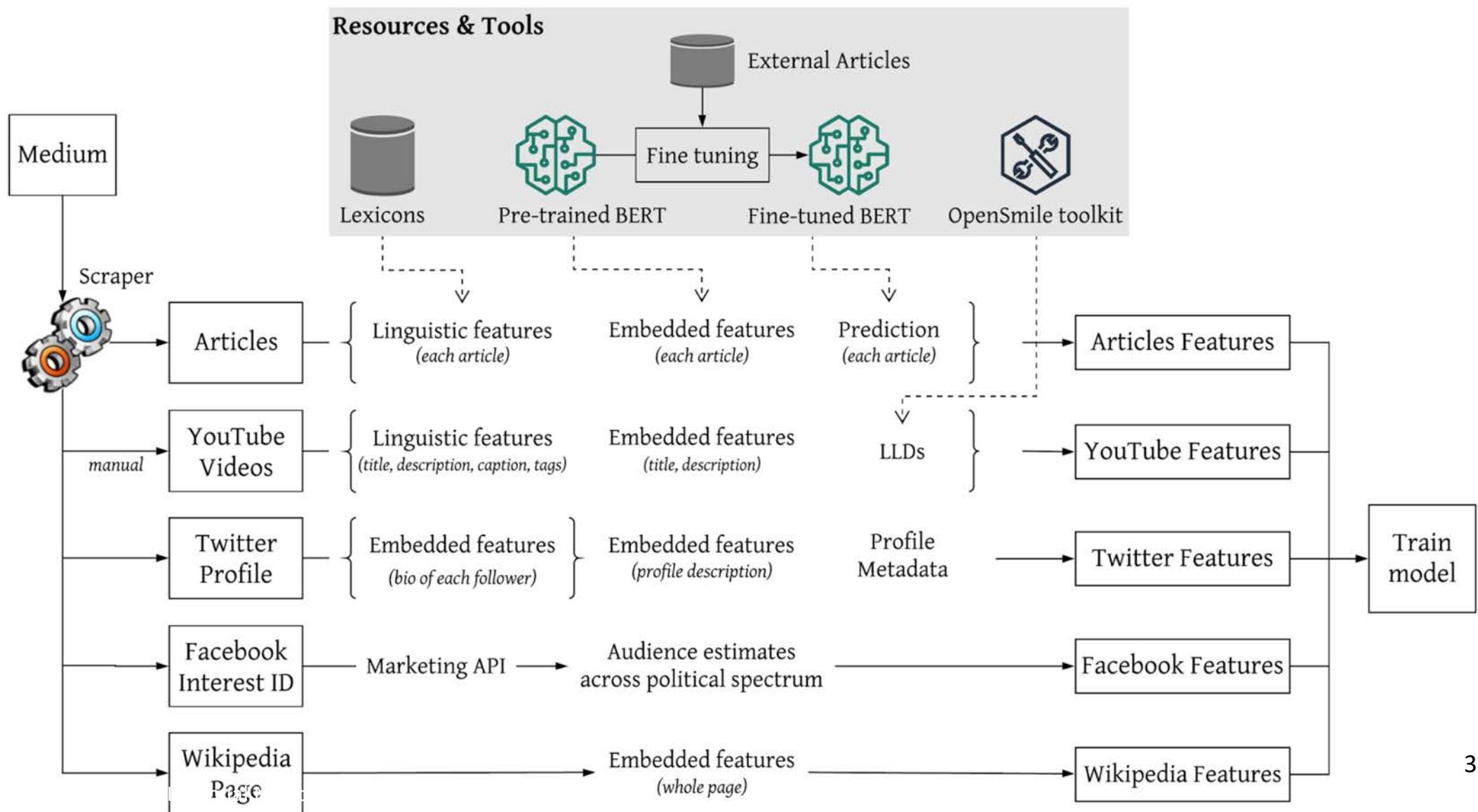
Become our patron  
on [patreon](#)

# Modeling Factuality and Bias

Name	URL	Bias	Factuality
London Web News	londonwebnews.com	Extreme Left	Low
Daily Mirror	www.mirror.co.uk	Left	Mixed
NBC News	www.nbcnews.com	Center-Left	High
Associated Press	apnews.com	Center	Very High
Gulf News	gulfnews.com	Center-Right	High
Russia Insider	russia-insider.com	Right	Mixed
Breitbart	www.breitbart.com	Extreme Right	Low

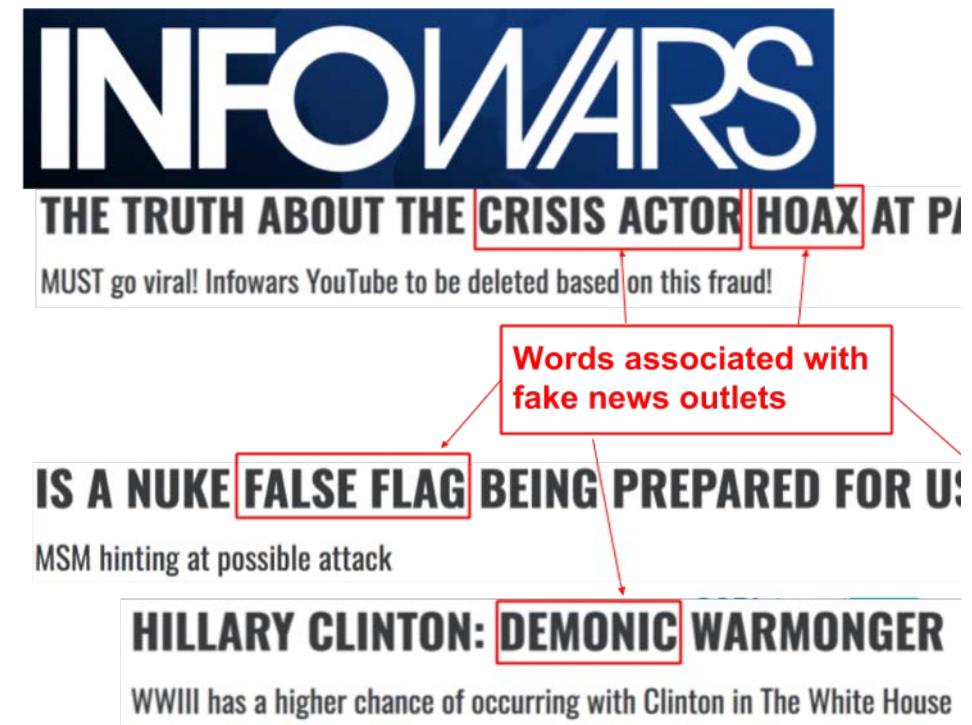
# Modeling Factuality and Bias in the News





# Article: Title & Body

- Connection between title vs. body
- Linguistic structure: function words, pronouns
- Length & complexity
- Sentiment
- Bias
- Subjectivity
- Topic
- Morality



# Wikipedia

- Has Wikipedia page?
- Embedding for the text in
  - Infobox
  - Summary
  - Content
  - Categories
  - Table of Contents

## *National Review*

From Wikipedia, the free encyclopedia

*This article is about the American magazine. For other uses, see National Review (disambiguation).*

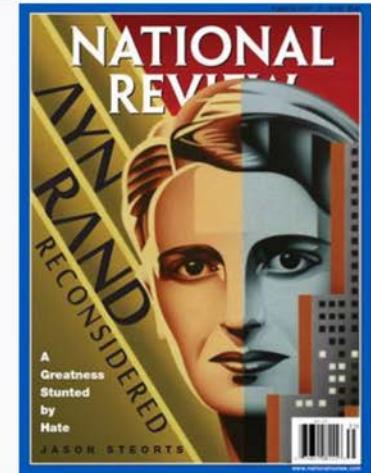
**National Review (NR)** is an American semi-monthly conservative editorial magazine focusing on news and commentary pieces on political, social, and cultural affairs. The magazine was founded by the author William F. Buckley Jr. in 1955.<sup>[3]</sup> It is currently edited by Rich Lowry.

Since its founding, the magazine has played a significant role in the development of conservatism in the United States, helping to define its boundaries<sup>[3]</sup> and promoting fusionism while establishing itself as a leading voice on the American right.<sup>[3][4][5]</sup>

The online version, **National Review Online**, is edited by Charles C. W. Cooke and includes free content and articles separate from the print edition.<sup>[6]</sup>

<b>Contents</b> [hide]
1 History
1.1 Background
1.2 Early years
1.3 Contributors
1.4 Mission to conservatives
1.5 Defining the boundaries of conservatism
1.6 After Goldwater
2 Political views and content
2.1 Donald Trump

**National Review**



National Review cover for August 30, 2010

<b>Editor</b>	Rich Lowry
<b>Categories</b>	Editorial magazine, conservatism
<b>Frequencies</b>	Bimonthly

# Twitter

- **Counts:** friends, statuses, favorites
- **Has Twitter account?**
- **Has location information?**
- **Verified?**
- **Years in existence**
- **Text description**



The image shows a Twitter profile for the user '@loki\_jcrusader' with the handle 'The Angry Patriot'. The profile picture features a bald eagle with its wings spread, set against a background of the American flag's stars and stripes. The bio reads: 'An Army Combat Vet ('99 to present). NCO. Crusader. Republican. Take back our country. #MAGA America First!'. The location is listed as 'USA' and the account was joined in February 2012. Below the profile, there are four tabs: 'Tweets' (645), 'Following' (404), 'Followers' (333), and 'Likes' (843). A 'Follow' button is visible. The main feed shows a single tweet from 'Dan Bongino' (@dbongino) dated July 22, 2022. The tweet reads: 'Dear Media Hacks, Now that the the FISA application has been released, & you've been completely humiliated for promoting the Russian collusion hoax, may I suggest that you all collectively submit letters of resignation? Your reputations are forever tarred. Thanks for considering.' The tweet has received 1.5K likes, 17K retweets, and 44K replies. At the bottom of the feed, it says 'The Angry Patriot Retweeted'.

# URL

- **Length**
- **Use of special characters**
  - digits
  - dashes
  - underscores
- **Use of https**
- **Is it hosted on a blog platform?**
- **Can the URL be chopped into meaningful words?**
- **What is the top-level domain?**

<http://www.angrypatriotmovement.com/>

<http://alternativemediasyndicate.com>

<http://100percentfedup.com/>

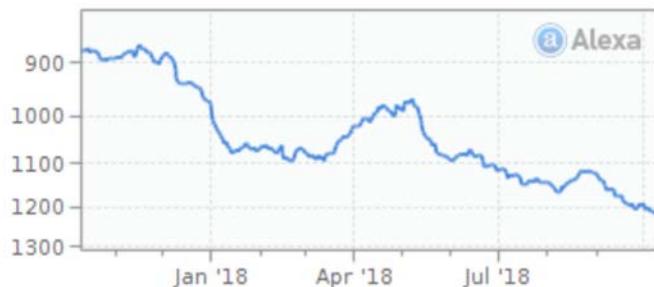
<http://abcnews.com.co>

# Web Traffic

How popular is abcnews.go.com?

## Alexa Traffic Ranks

How is this site ranked relative to other sites?



Global Rank ?

1,220 ▼ 92

Rank in United States ?

324

How popular is abcnews.com.co?

## Alexa Traffic Ranks

How is this site ranked relative to other sites?



Global Rank ?

2,767,435 ▼ 664,690

# Youtube

<https://youtube-politics.herokuapp.com/media/the-patriot-post>

YouTube Politics Media Annotate About Search...

## The Patriot Post



### Breitbart's Ben Shapiro Gives CNN's Piers Morgan a Smackdown

COMPLETE Ben Shapiro CUTS Piers Morgan DOWN

YouTube

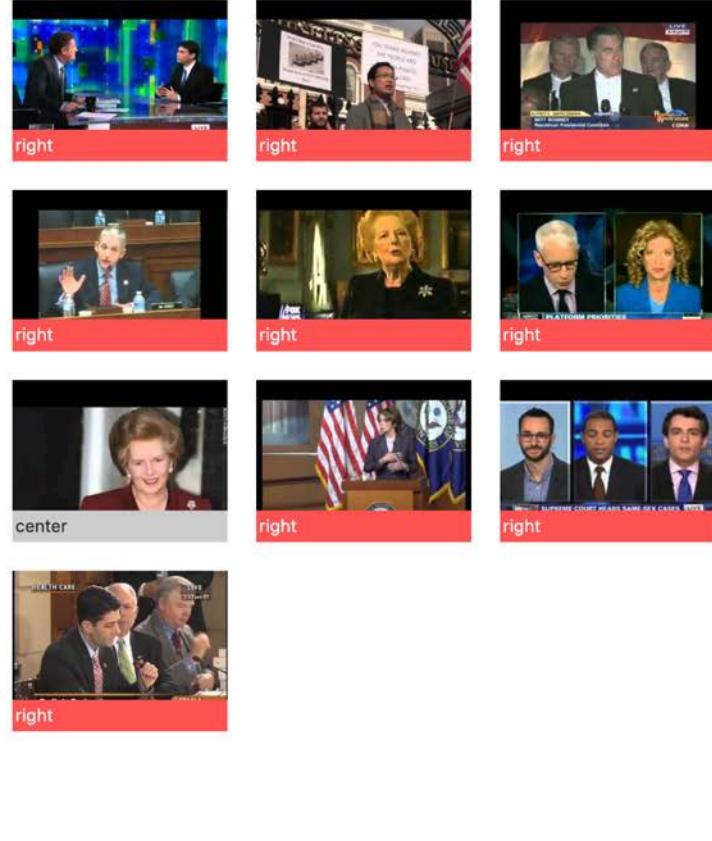
#### Bias distribution



#### Video stats

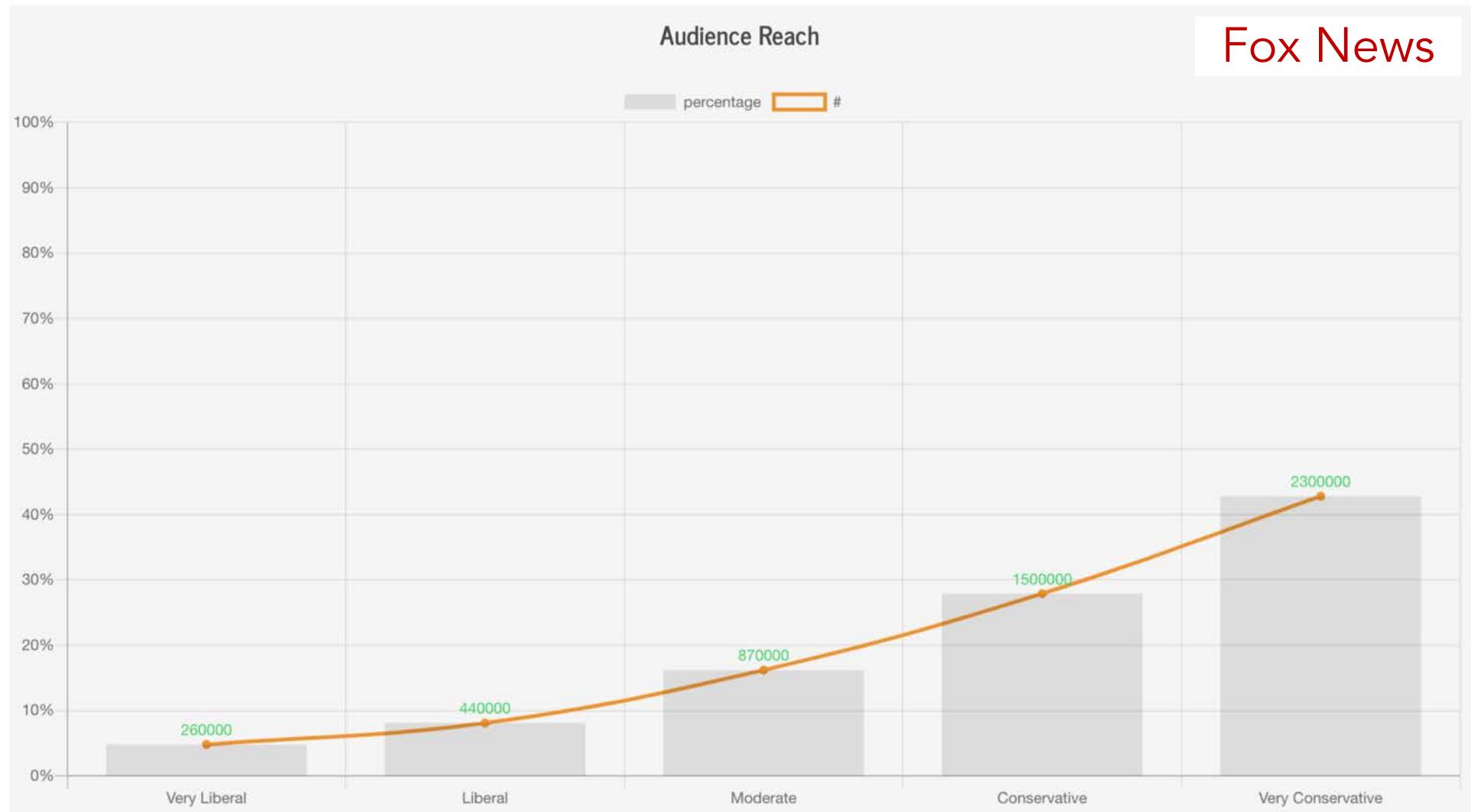
Views	2087
Likes	35
Dislikes	0
Comments	14

#### Speech episodes



right right right right right right center right right right

# Audience Reach (in Facebook)



# Audience Bias (in Twitter)

RT

Twitter User Based Bias Prediction



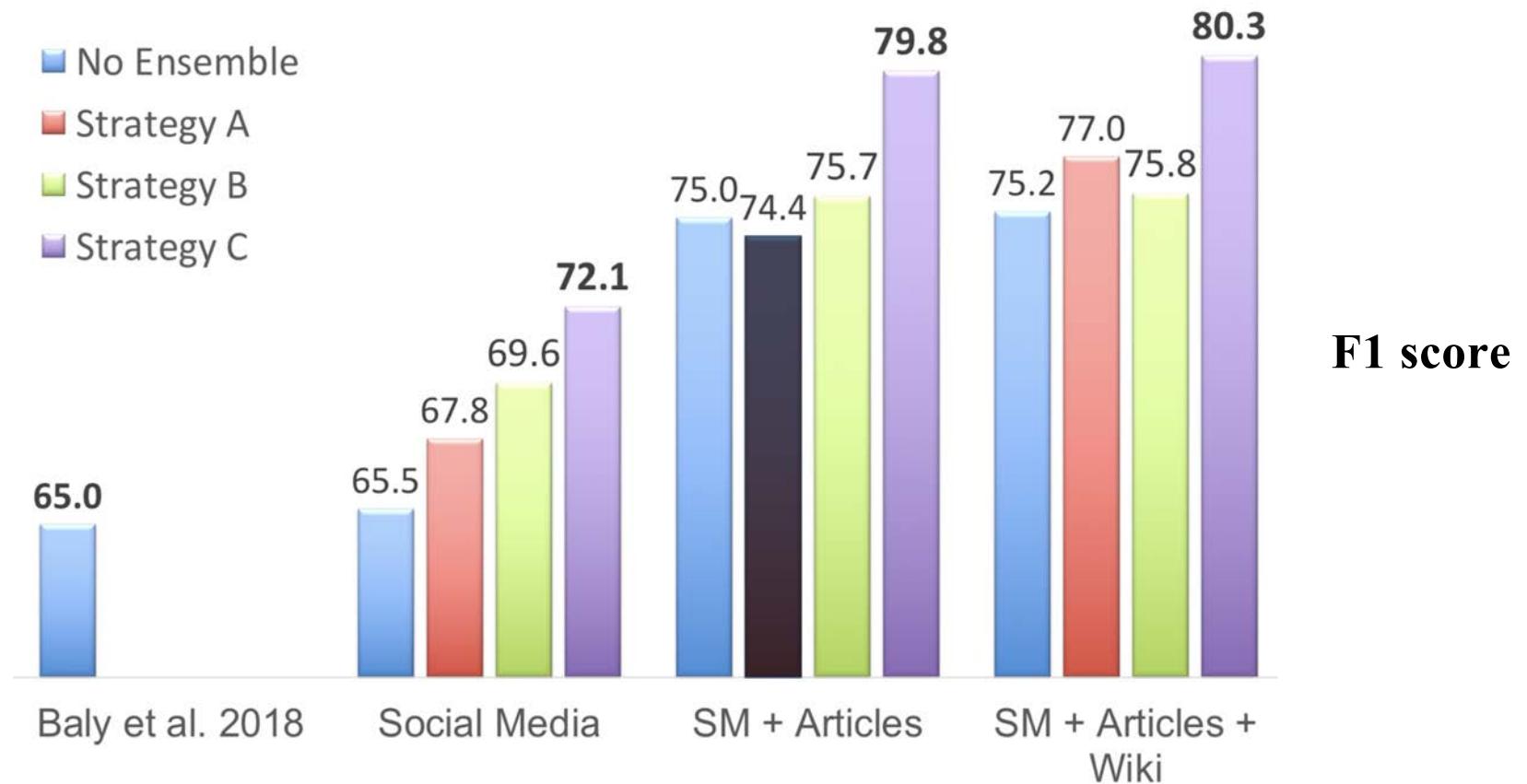
# Results

F1 score

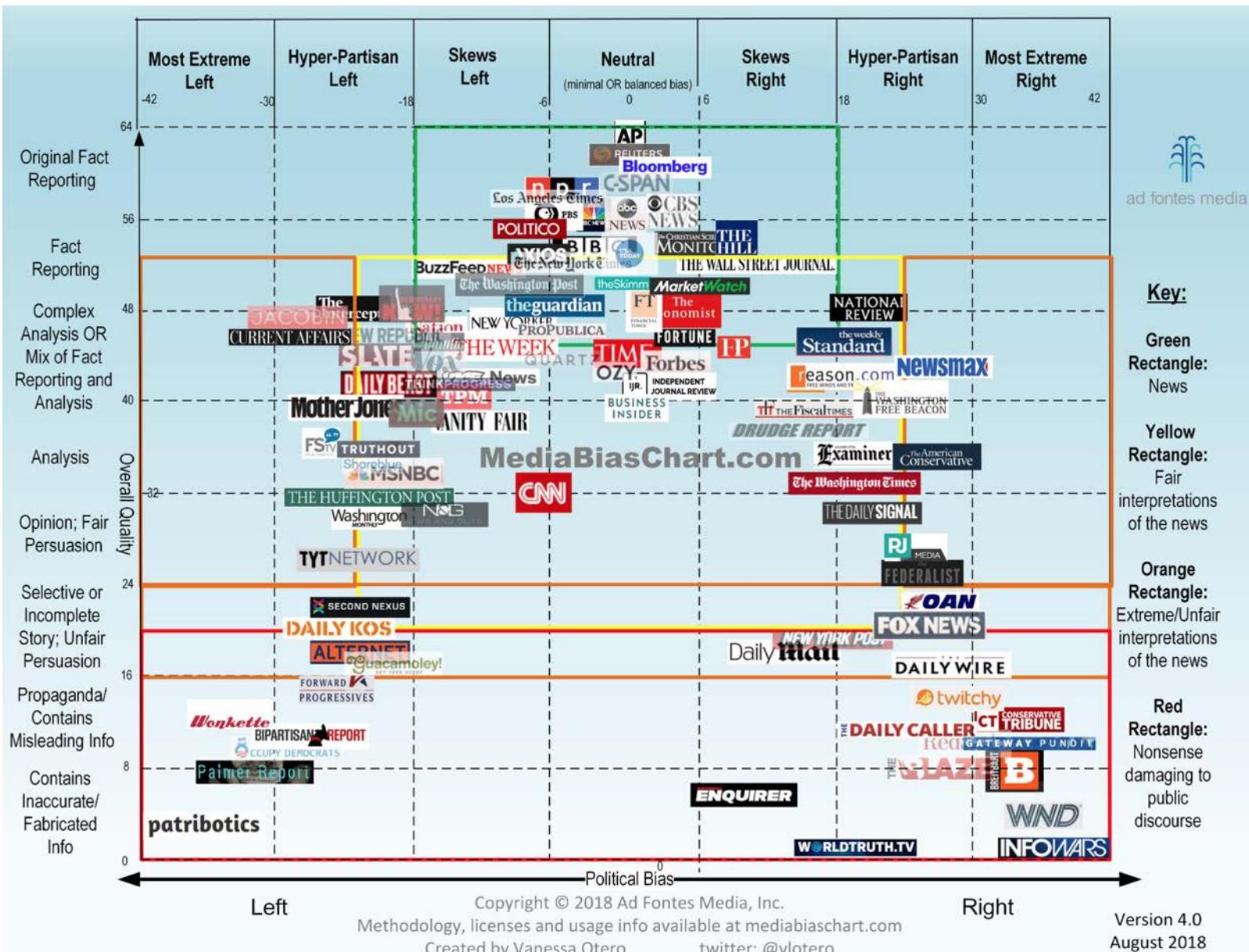
	Features	Dim.	Bias	Factuality
1	Majority class baseline	–	24.5	21.0
2	(Baly et al. 2018)	764	65.0	62.2
3	Articles (NELA)	141	63.3	56.6
4	Articles (BERT)	768	72.7	63.5
5	Articles (Agg. Predictions)	3	34.3	53.9
6	Articles (NELA + BERT + Agg. Predictions)	912	75.7	79.8
7	Wikipedia (BERT)	768	60.7	49.6
8	Twitter (Media Profiles)	779	51.6	48.2
9	Twitter (Followers)	768	70.6	55.5
11	YouTube (Metadata + Text + Acoustics).	2,186	43.1	34.3
12	Facebook (Audience ideology)	6	22.9	23.9
13	All Social Media (lines 8–12)	3,739	65.5	56.5
14	All sources (lines 6–12)	5,419	75.2	77.6

(older work) Ramy Baly, Georgi Karadzhov, Dimitar Alexandrov, James R. Glass, Preslav Nakov:  
*Predicting Factuality of Reporting and Bias of News Media Sources*. EMNLP 2018

# Media Bias: *left-center-right*

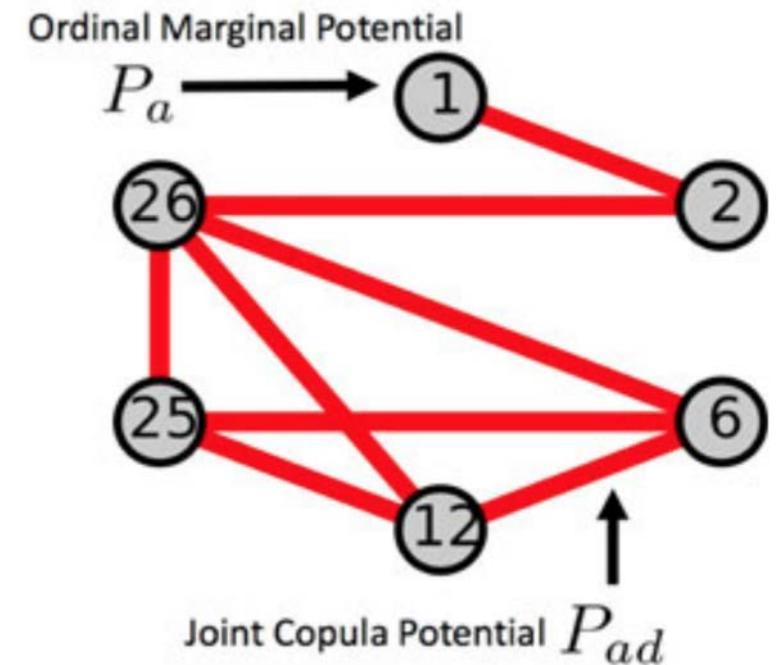


(older work) Ramy Baly, Georgi Karadzhov, Dimitar Alexandrov, James R. Glass, Preslav Nakov:  
*Predicting Factuality of Reporting and Bias of News Media Sources*. EMNLP 2018



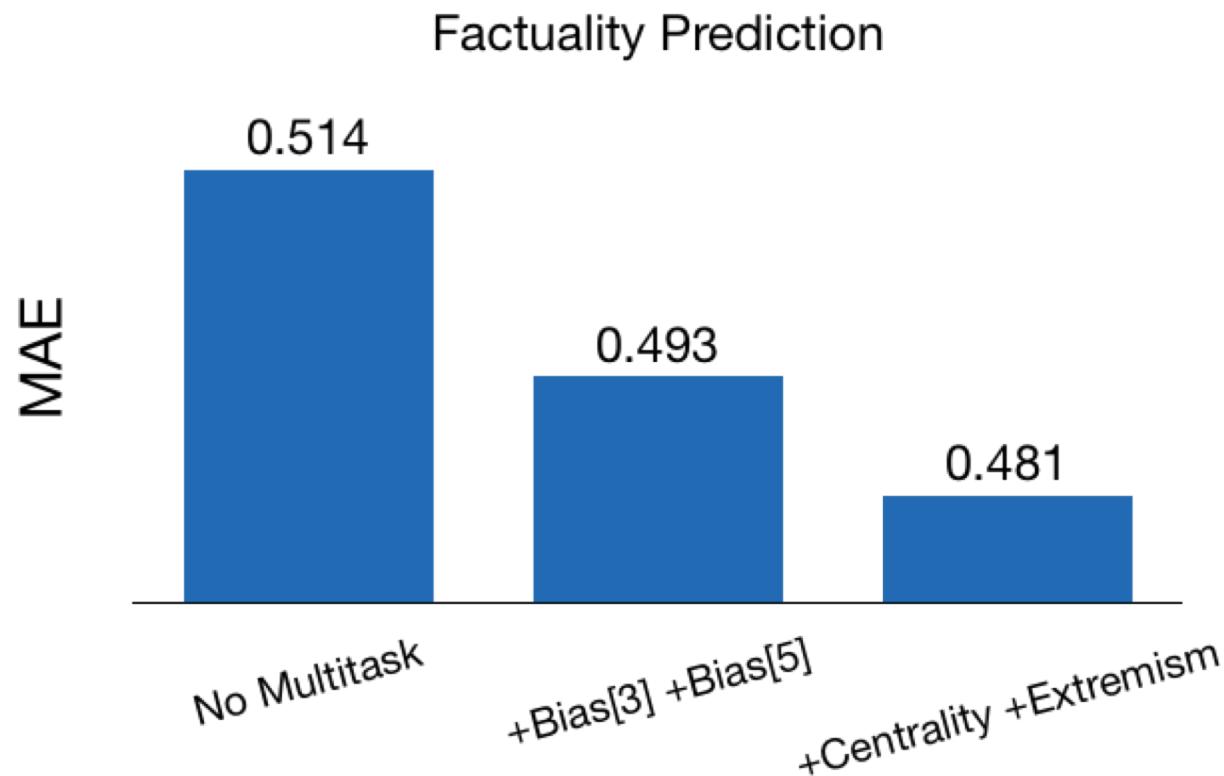
# Multitask Ordinal Regression for Factuality & Bias

- **Ordinal regression**
- **Learn jointly**
  - Factuality
  - Left-center-right bias
    - 7-point
    - 5-point
    - 3-point
  - Centrality
  - Hyper-partisanship



Ramy Baly, Georgi Karadzhov, Abdelrhman Saleh, James R. Glass, Preslav Nakov:  
*Multi-Task Ordinal Regression for Jointly Predicting the Trustworthiness and the Leading Political Ideology of News Media.*  
NAACL-HLT 2019

# Multitask Ordinal Regression for Factuality



Ramy Baly, Georgi Karadzhov, Abdelrhman Saleh, James R. Glass, Preslav Nakov:  
*Multi-Task Ordinal Regression for Jointly Predicting the Trustworthiness and the Leading Political Ideology of News Media.*  
NAACL-HLT 2019

# Propaganda

# Propaganda

- “Expression **deliberately** designed to **influence** the opinions/actions of other individuals or groups with reference to predetermined ends.”

Institute for Propaganda Analysis

# Proppy

<http://proppy.qcri.org>

The screenshot shows the Proppy web application interface. At the top center is the Proppy logo, which consists of the word "PROPPY" in white capital letters inside a blue circular emblem with a grid pattern. Below the logo is the tagline "The propaganda analyzer". On the left side, there is a vertical list titled "List of the last 24 hours' events:" containing seven news items. On the right side, there is a horizontal bar chart titled "estimated amount of propagandistic content" ranging from 0.0 to 1.0. The chart has six segments, with the fourth segment being the longest and colored dark blue.

List of the last 24 hours' events:

- Malaysian ex-PM Najib hit with 25 new charges over 1MDB scandal
- Air pressure error causes mass bleeding on India flight
- North Dakota man rescues animals from Hurricane Florence
- Japan eager to get in on rise of 'flying cars'
- Abbas to hold meeting with ministers, diplomats in New York
- Trump blames OPEC for rising oil prices, says U.S. 'will remember'
- New James Bond movie to be directed by Cary Joji Fukunaga

estimated amount of propagandistic content

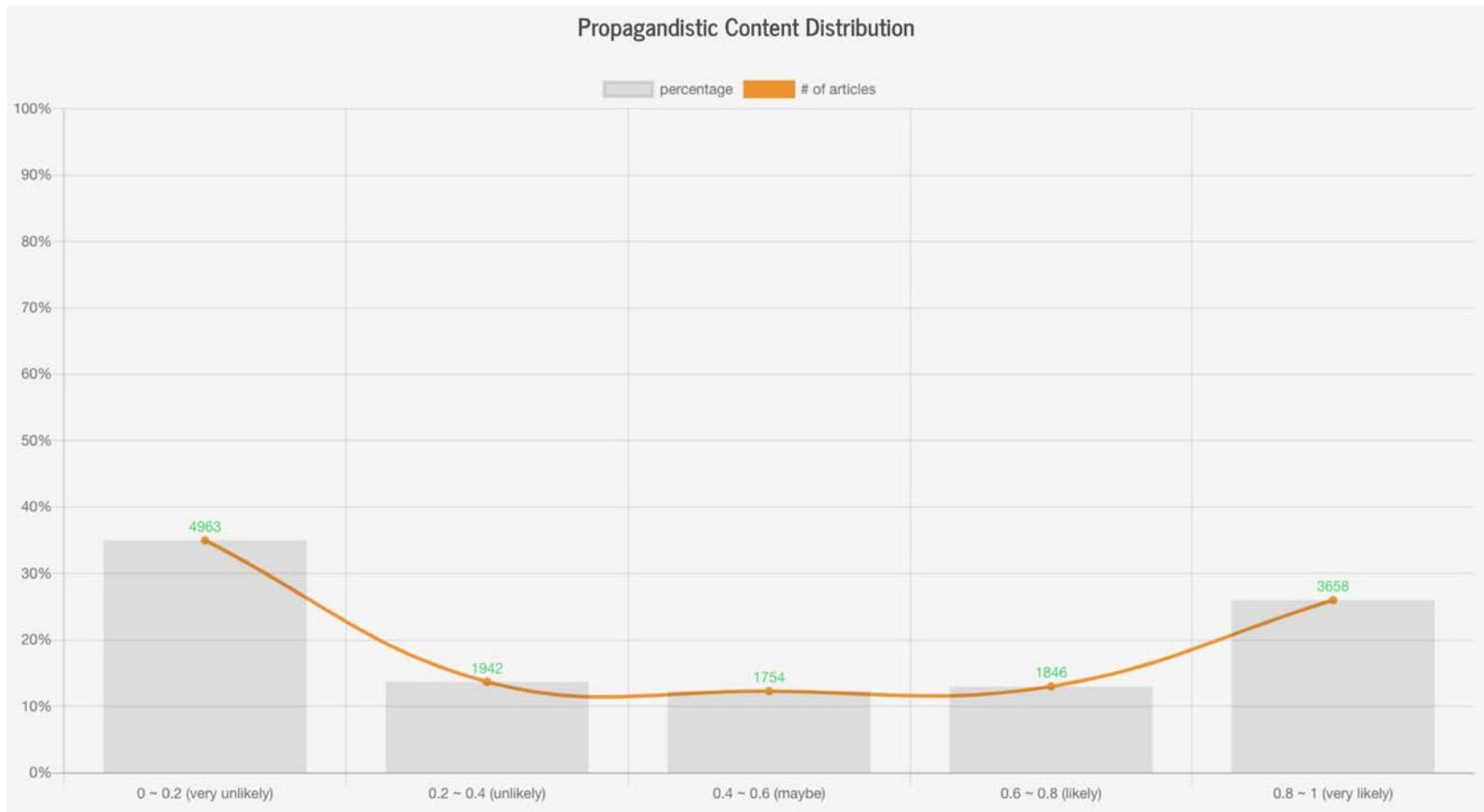
**Malaysian ex-PM Najib hit with 25 new charges over 1MDB scandal**  
Kuala Lumpur, Kuala Lumpur, Malaysia --2018-09-20  
Malaysia's former Prime Minister Najib Razak has been hit with 25 new corruption charges linked to the 1MDB scandal. The new charges include four counts of abuse of power and 21 counts of money laundering over hundreds of millions of dollars of fund..  
Source:[www.aljazeera.com](http://www.aljazeera.com)

**Najib faces 25 charges linked to \$681m in 1MDB case | Bangkok Post: news**  
British Virgin Islands --2018-09-13  
Former Malaysian prime minister Najib Razak (centre) arrives at the High Court in Kuala Lumpur on Thursday. (EPA-EFE photo) KULA LUMPUR: Former Malaysian premier Najib Razak seeks trial for charges linked to \$681 million that appeared in his persona..  
Source:[www.bangkokpost.com](http://www.bangkokpost.com)

Alberto Barrón-Cedeño, Giovanni Da San Martino, Israa Jaradat, Preslav Nakov:  
*Proppy: A System to Unmask Propaganda in Online News.* AAAI 2019: 9847-9848



<http://www.tanbih.org>



Yifan Zhang, Giovanni Da San Martino, Alberto Barrón-Cedeño, Salvatore Romeo, Jisun An, Haewoon Kwak, Todor Staykovski, Israa Jaradat, Georgi Karadzhov, Ramy Baly, Kareem Darwish, James Glass, Preslav Nakov.  
*Tanbih: Get To Know What You Are Reading.* EMNLP 2019

<http://www.tanbih.org>



Mail  
Online

Ex-CIA agent convicted in Italy for her role in the  
abduction of a Muslim cleric flees to the US

Propaganda? likely

MILITARY

1 of 2

Yifan Zhang, Giovanni Da San Martino, Alberto Barrón-Cedeño, Salvatore Romeo, Jisun An, Haewoon Kwak, Todor Staykovski, Israa Jaradat, Georgi Karadzhov, Ramy Baly, Kareem Darwish, James Glass, Preslav Nakov.  
*Tanbih: Get To Know What You Are Reading.* EMNLP 2019

53



The  
Guardian

G

Sparks fly after roadworks leave power pole in middle of  
Newcastle street

AUSTRALIA NEWS OCT

2 of 2

# Propaganda Techniques



Name Calling



Band Wagon

reductio ad Hitlerum  
flag-waving  
minimisation  
exaggeration  
black-and-white fallacy  
appeal to prejudice

thought-terminating cliches  
whataboutism  
bandwagon  
causal oversimplification  
appeal to authority  
intentional vagueness  
cognitive dissonance  
name calling  
obfuscation  
loaded language

# Fine-Grained Propaganda Detection

Stereotyping\_name\_calling\_or\_labeling

- 1 Manchin says Democrats acted like babies at the SOTU
  - 2 Democrat West Virginia Sen. Joe Manchin says his colleagues' refusal to stand or applaud during President Donald Trump's State of the Union speech was disrespectful and a signal that
- Black-and-white\_Fallacy
- the party is more concerned with obstruction than it is with progress.

Loaded\_language

- 4 In a glaring sign of just how stupid and petty things have become in Washington these days, Manchin was invited on Fox News Tuesday morning to discuss how he was one of the only Democrats in the chamber for the State of the Union speech

Exaggeration

Loaded\_language

not looking as though Trump killed his grandma.

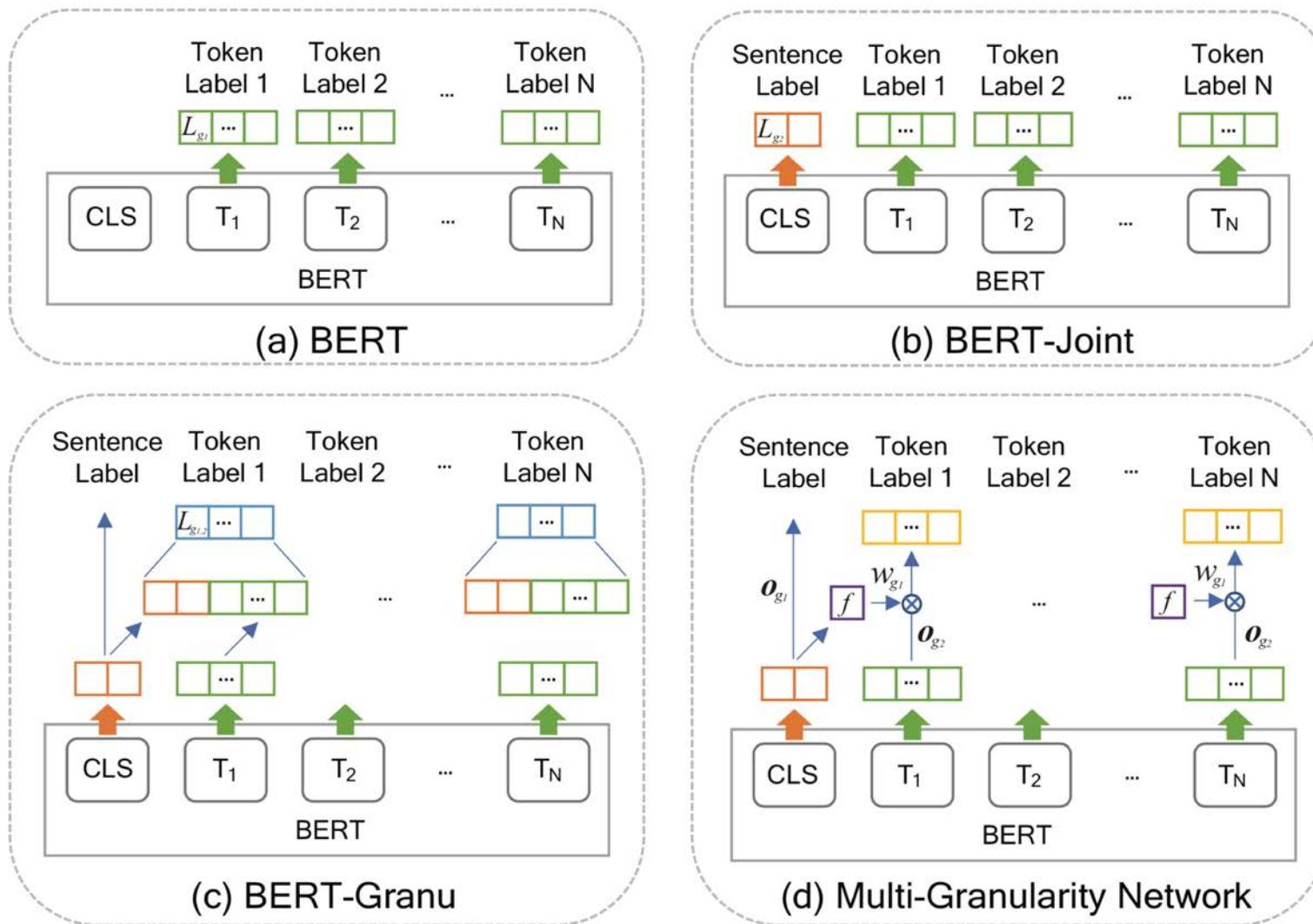
- 6 As Manchin noted, many Democrats bolted as soon as Trump's speech ended in an apparent effort to signal

Exaggeration

they can't even stomach being in the same room as the president

## New Dataset

- 18 techniques
- 350k words
- 400 man hours
- 7.3k instances



Giovanni Da San Martino, Seunghak Yu, Alberto Barrón-Cedeño, Rostislav Petrov, Preslav Nakov  
*Fine-Grained Analysis of Propaganda in News Articles*. EMNLP 2019

## Results: Fragment-Level

Model	Spans			Full Task		
	P	R	F <sub>1</sub>	P	R	F <sub>1</sub>
BERT	39.57	36.42	37.90	21.48	<b>21.39</b>	21.39
Joint	39.26	35.48	37.25	20.11	19.74	19.92
Granu	43.08	33.98	37.93	23.85	20.14	21.80
Multi-Granularity						
ReLU	43.29	34.74	38.28	23.98	20.33	21.82
Sigmoid	<b>44.12</b>	<b>35.01</b>	<b>38.98</b>	<b>24.42</b>	21.05	<b>22.58</b>

# Results: Sentence-Level

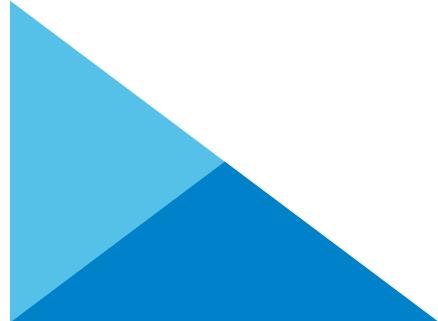
Model	Precision	Recall	F1
All-Propaganda	23.92	1.00	38.61
BERT	<b>63.20</b>	53.16	57.74
BERT-Granu	62.80	55.24	58.76
BERT-Joint	62.84	55.46	58.91
MGN Sigmoid	62.27	59.56	60.71
MGN ReLU	60.41	<b>61.58</b>	<b>60.98</b>

# Shared Tasks

The image displays three separate web pages related to shared tasks:

- HACK THE NEWS DATATHON**: A dark-themed website featuring a red banner with white text "HACK THE NEWS DATATHON". Below it is a white button with the text "LET'S HACK NEWS MANIPULATION!". The navigation bar includes links for "INTERNATIONAL DATA SCIENCE SOCIETY", "TASK", "RULES/DATES", and "REGISTER". At the bottom, there are links for "WHY?", "PARTICIPATE", "HOW IT WAS BORN", "THE CASE", and "THE JURY", along with a "LEARN MORE" button.
- SHARED TASK ON FINE-GRAINED PROPAGANDA DETECTION @NLP4IF 2019**: A website with a background image of a city skyline at night. The main title "SHARED TASK ON FINE-GRAINED PROPAGANDA DETECTION" is displayed in large white letters, followed by "@NLP4IF 2019" in a slightly smaller font.
- SEMEVAL 2020 TASK 11**: A website with a dark teal background. The title "SEMEVAL 2020 TASK 11" is at the top, followed by the subtitle "DETECTION OF PROPAGANDA TECHNIQUES IN NEWS ARTICLES" in a larger font.

# AI for Regulatory Practice: Application to Media





# Media-Telecom Catalyst Programme





TEC  
Mahindra

AP

12.20.18.12.2020 analysis\_0001

Earlier Seoul, South Korea

STATE OF THE UNION, CNN

CNN SOTU

LIVE

TRUMP AND KIM HELD PRIVATE TALKS FOR NEARLY AN HOUR

CNN

TRUMP & EXECUTIVE ORDER BANNING U.S. FIRMS FROM PURCHASING OR TRANSFER

Metadata Contextual Metadata Regulatory Compliance Report

Sentiments

- Yellow: Neutral
- Green: Positive
- Red: Negative

Bias

Claims Worth

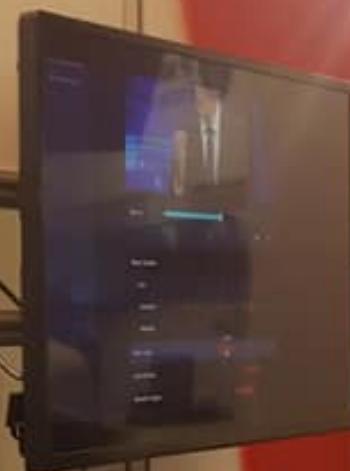
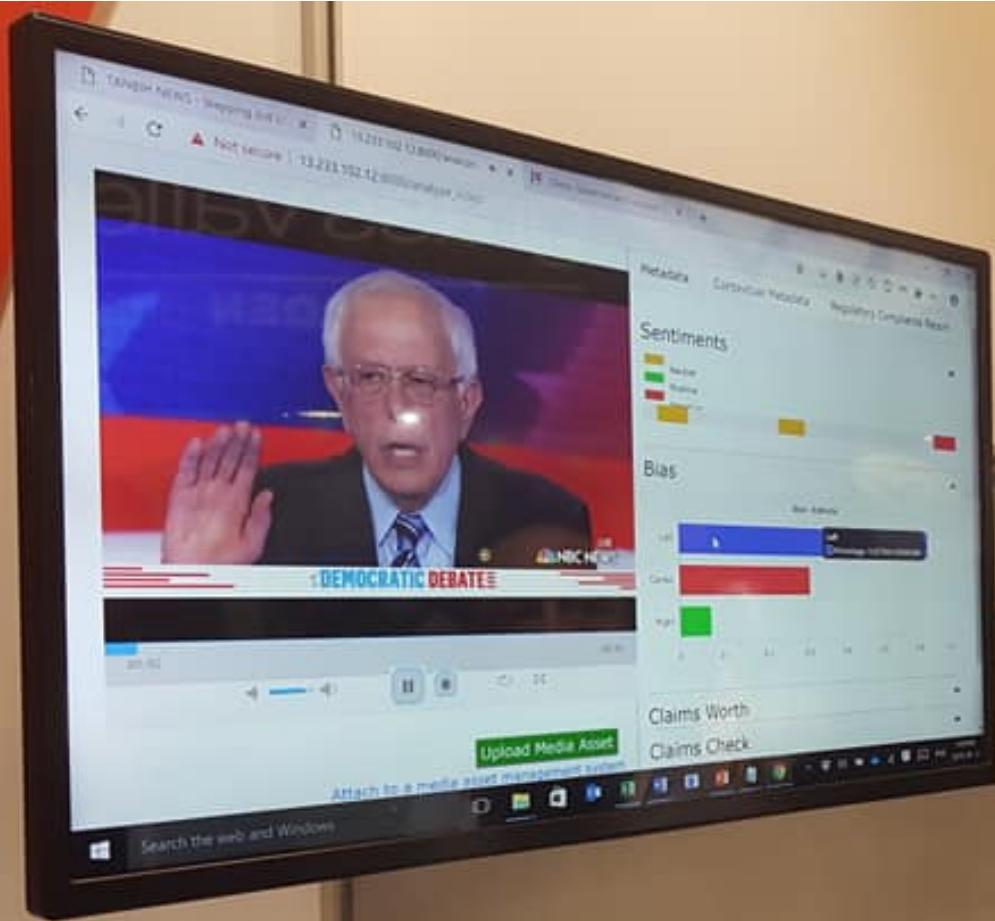
Claims Check

- "The Obama administration was begging for a meeting" with Kim Jong Un. X
- "The governor having a meeting ... isn't covered by the open meetings act." ✓

Upload Media Asset

Attach to a media asset management system

Search for news and videos



**BREAKING NEWS**

JUAN G

Smart Tourism Experience



tmforum

RTE AR  
2023



← Tanbih v 0.6.0 v

Export v



Info



Tags



Search



default ^

POST /article/propaganda

GET /article/propaganda

POST /article/propaganda/sentences

GET /article/propaganda/sentences

POST /article/bias

GET /article/bias

POST /article/checkworthiness

GET /article/checkworthiness

POST /claim/check

GET /claim/check

```
1 # Example YAML to get you started quickly.
2 # Be aware that YAML has indentation based coping. Read Only
3 # Code completion support is available so start typing for
   available options.
4 swagger: '2.0'
5
6 # This is your document metadata
7 info:
8   version: "0.6.0"
9   title: Tanbih API
10  description: |
11    Tanbih API allows user to estimate propaganda score on
      both article level
12    and sentences level and to predict bias on given piece
      of text. With current
13    version, the article level propaganda estimation is
      implemented.
14  Requests for estimation is done in a asynchronical way.
      Client will POST
15  a request for estimation, when successful, get a key in
      return; then client
16  will use this key in query parameter to pull the
      estimation result from
      server.
17
18
19 host: webapi.tanbih.org
20
21 schemes:
22   - http
```

Last Saved: 4:36:57 pm - Sep 9, 2019

✓ VALID ▾

# Tanbih API

0.6.0

[ Base URL: webapi.tanbih.org ]

Tanbih API allows user to estimate propaganda score on both article level and sentences level and to predict bias on given piece of text. With current version, the article level propaganda estimation is implemented. Requests for estimation is done in a asynchronical way. Client will POST a request for estimation, when successful, get a key in return; then client will use this key in query parameter to pull the estimation result from server.

## Schemes

HTTP ▾

## default

**POST** /article/propaganda **GET** /article/propaganda

(en) to join them to call market had not been globally certified and reason by company.

/ office

# QCRI team honoured for innovation at tech expo

TRIBUNE NEWS NETWORK

DOHA

A team of scientists from Qatar Computing Research Institute (QCRI) at Hamad Bin Khalifa University have received recognition for their contribution to a media-telecom catalyst project on AI indexing for regulatory compliance that uses artificial intelligence to identify, index and measure political balance.

The project aims to detect false claims and to enhance the accuracy of video content.

Developed in partnership with Al Jazeera, Associated Press, RTE Ireland, Tech Mahindra, V-Nova and Metaliquid, the project seeks to support broadcasters and content owners with their regulatory measures and compliance processes.

QCRI worked on the com-



A team of scientists from Qatar Computing in Amsterdam.

pliance use of the project, developing the fact checking, news coverage, political bias detection and graphic content detection mechanism.

The team of QCRI scientists was honoured with an award at IBC2019, a leading media, entertainment and technology

conference held in Amsterdam in September.

The global gathering saw more than 50,000 industry experts and researchers from 150 countries come together to showcase their latest innovations and participate in a number of topical discussions.



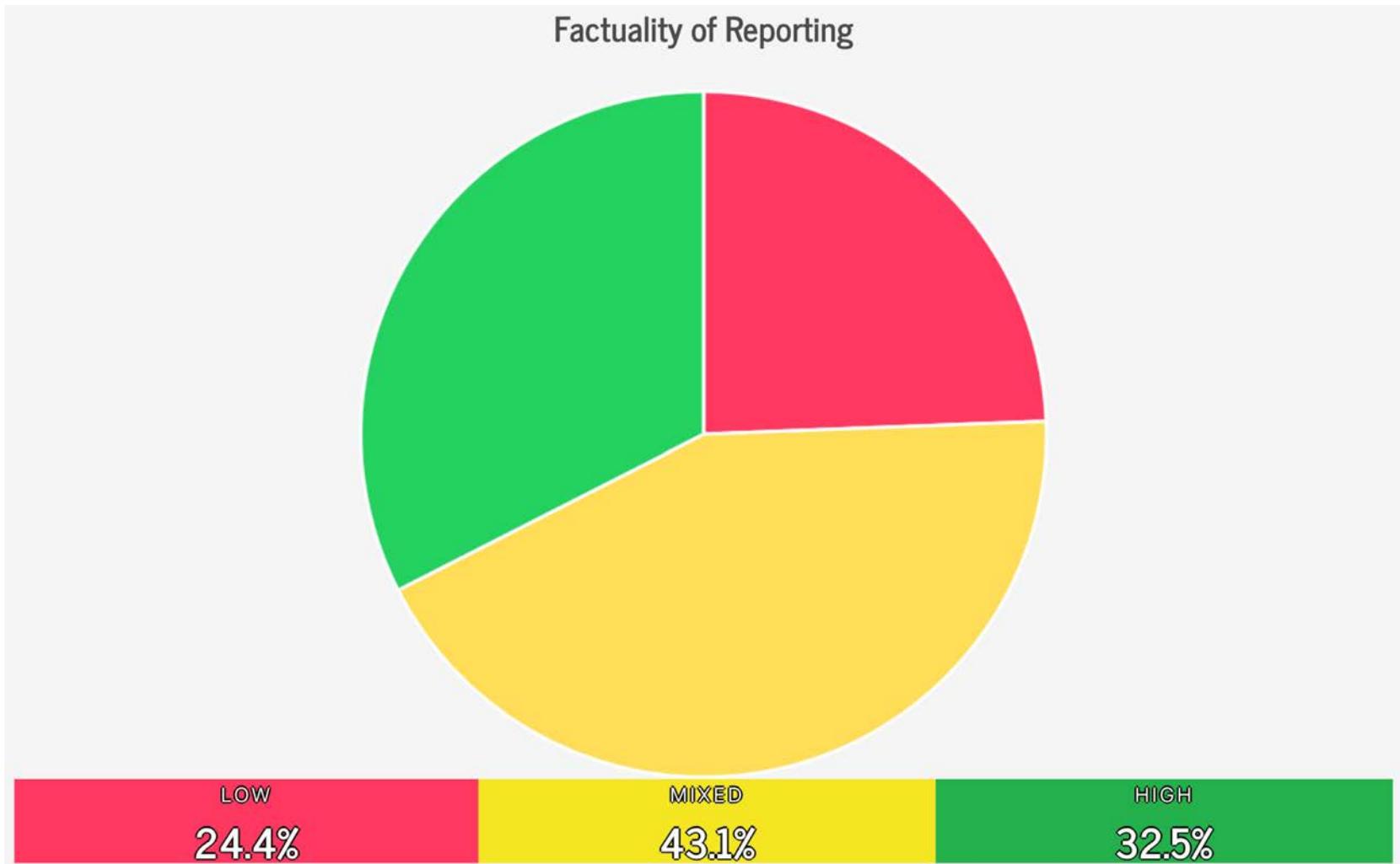
<http://www.qatartribune.com/news-details/id/173546/qcri-team-honoured-for-innovation-at-tech-expo>

# Demo

Try Tanbih: <http://www.tanbih.org>

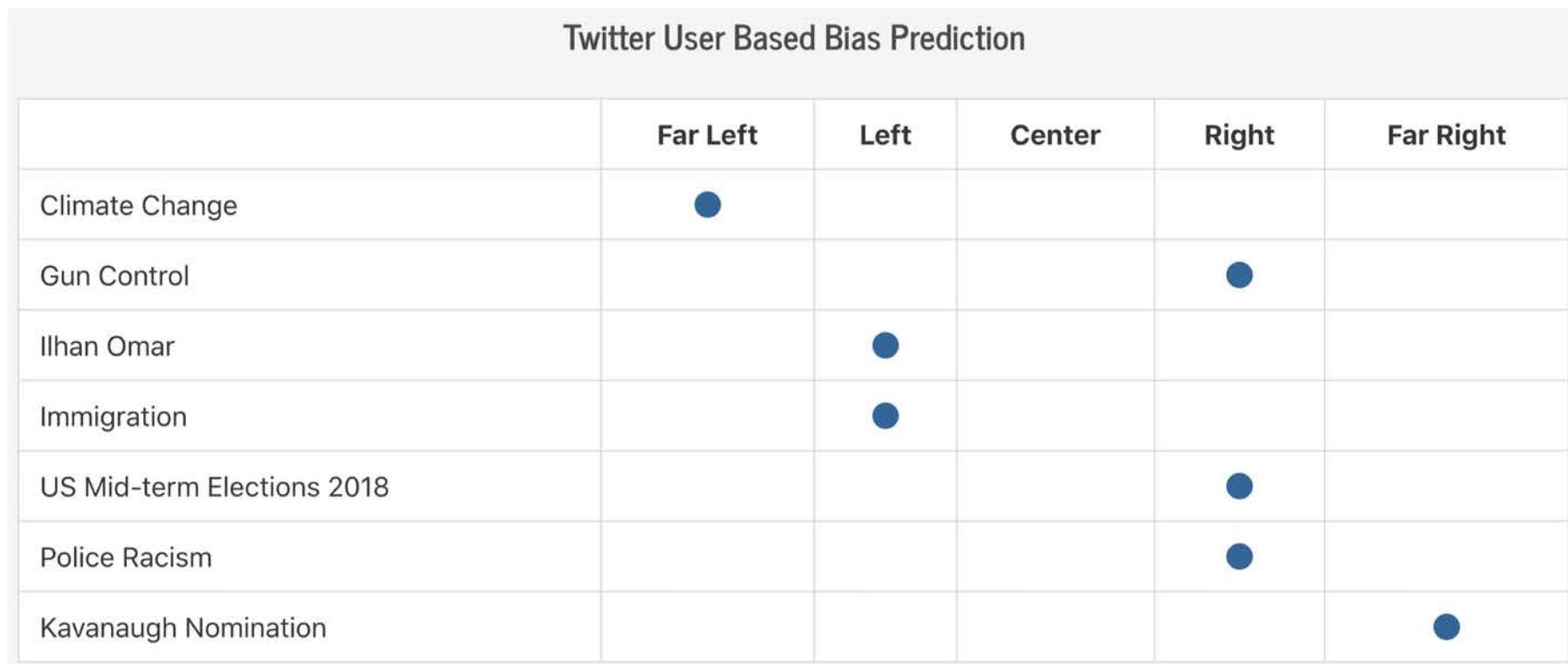


<http://www.tanbih.org>





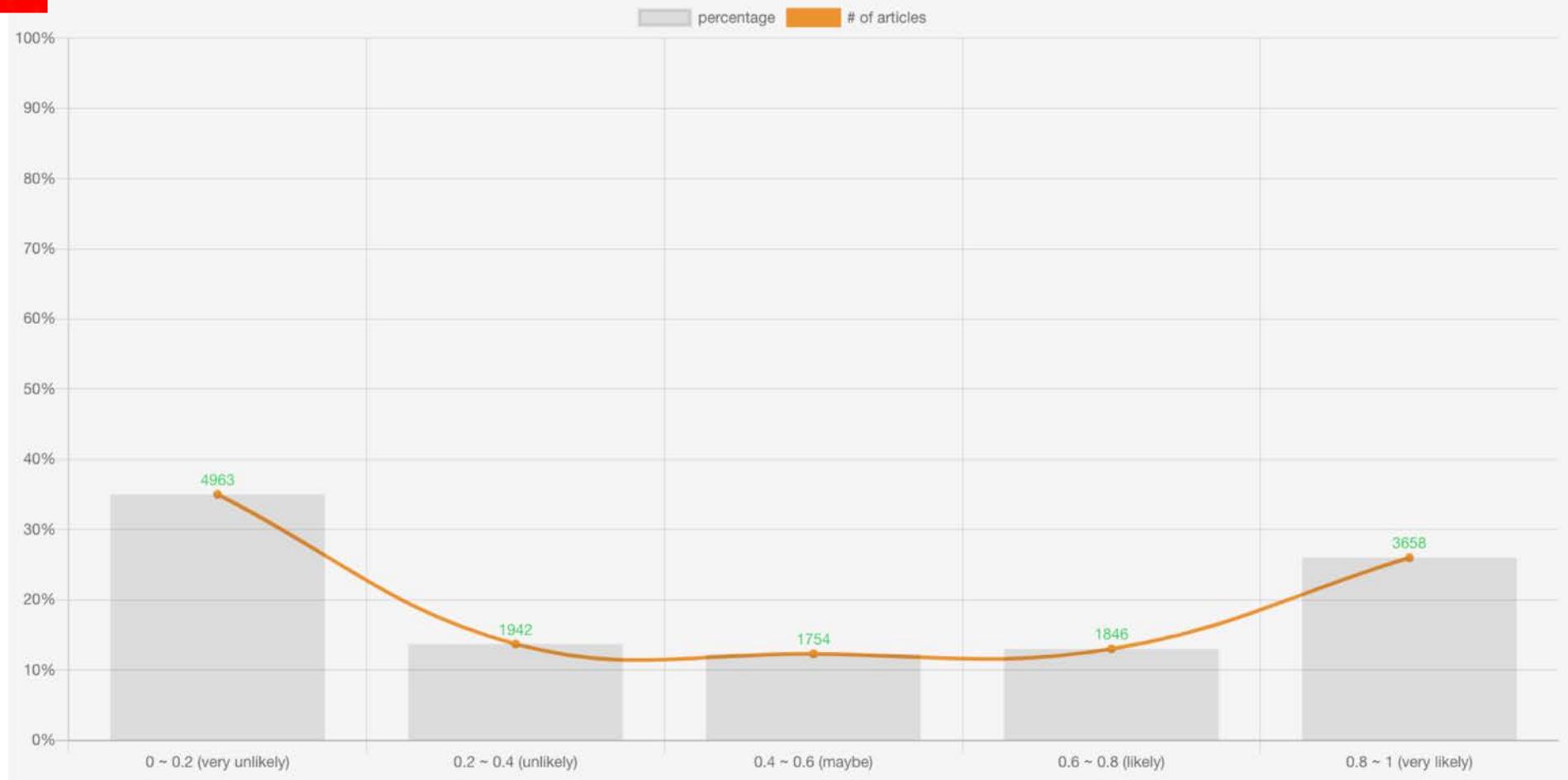
<http://www.tanbih.org>





<http://www.tanbih.org>

### Propagandistic Content Distribution



<http://www.tanbih.org>



Centrality



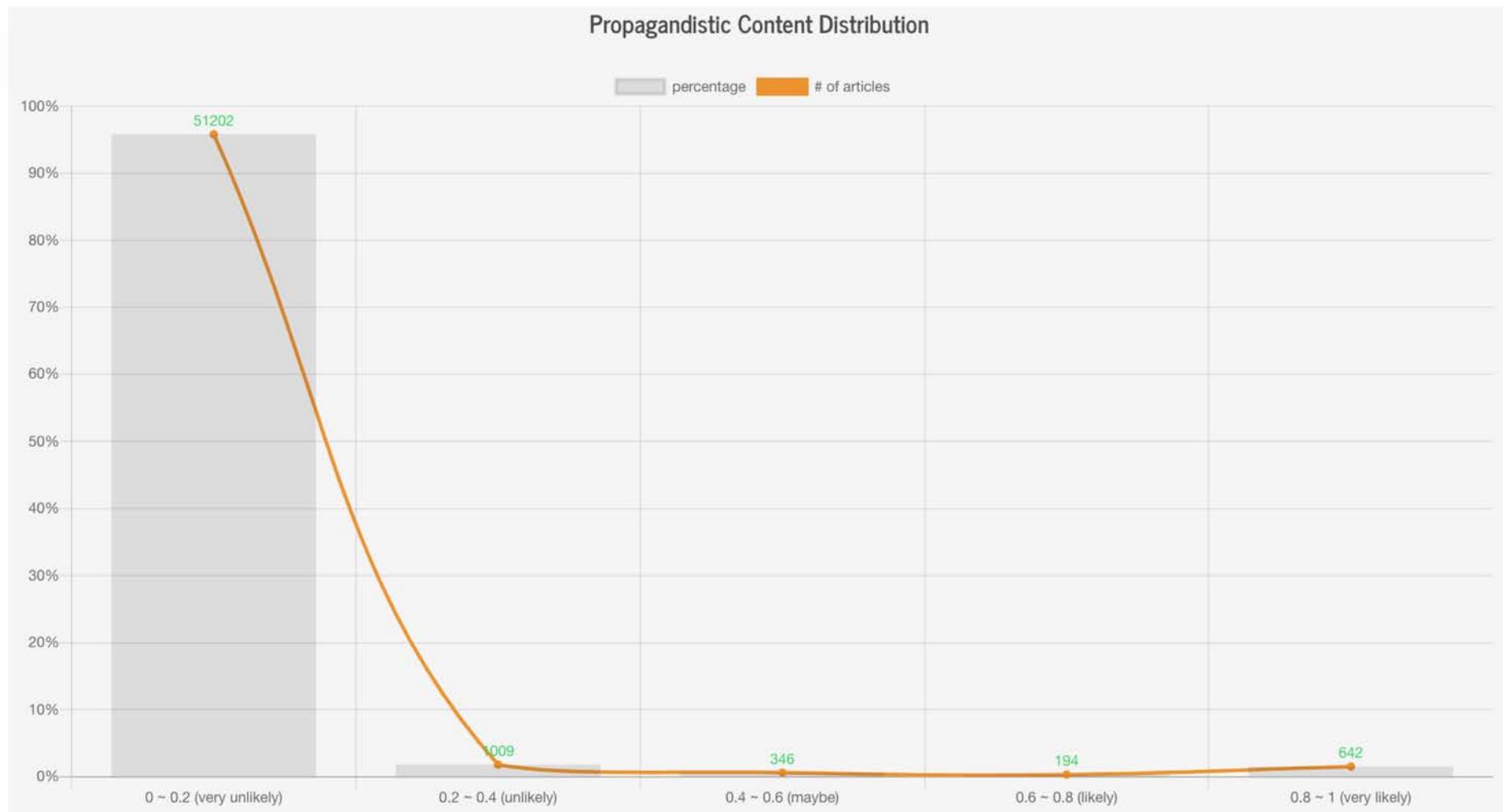
57%

Hyper-Partisanship



0%

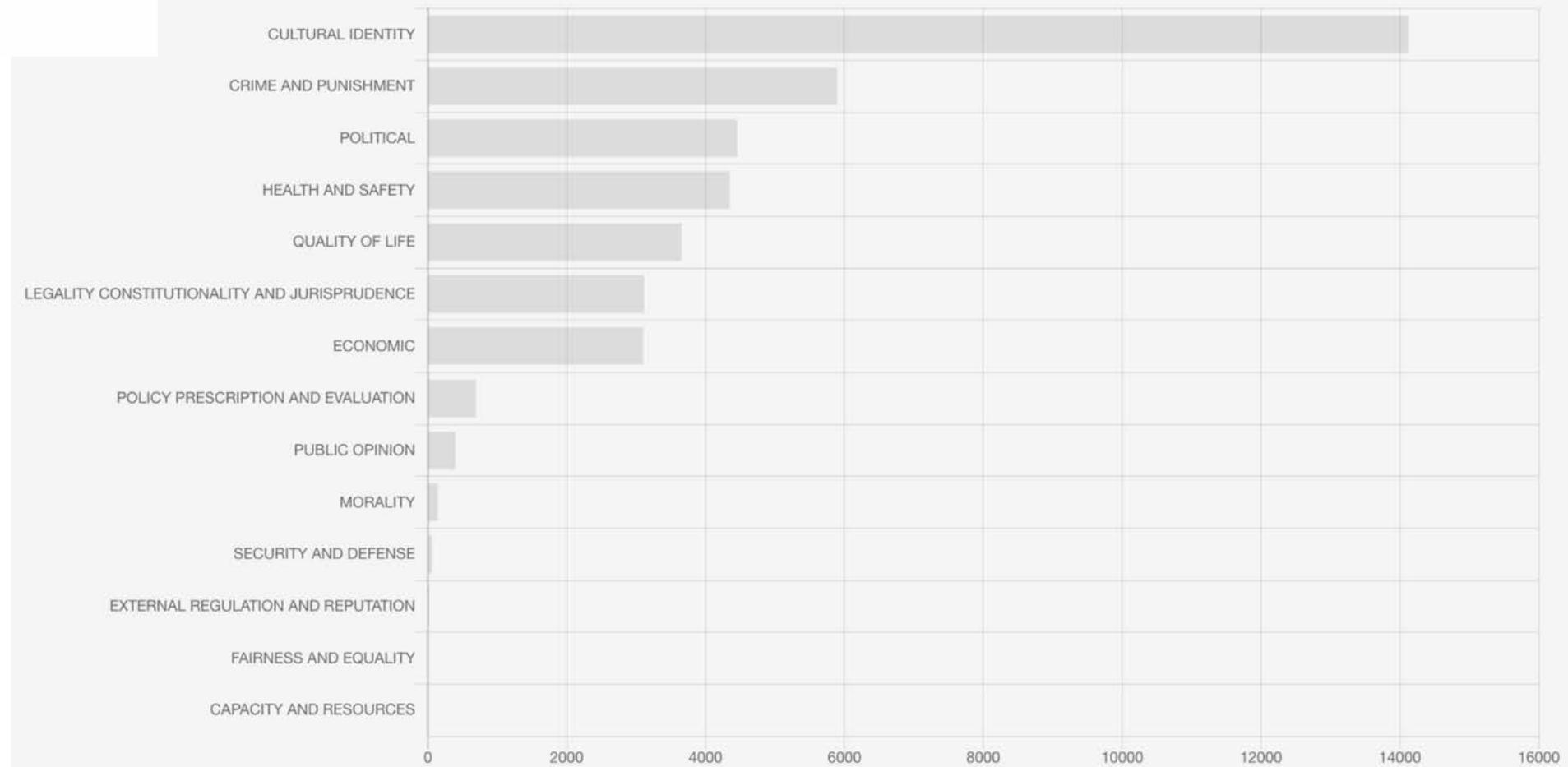
<http://www.tanbih.org>



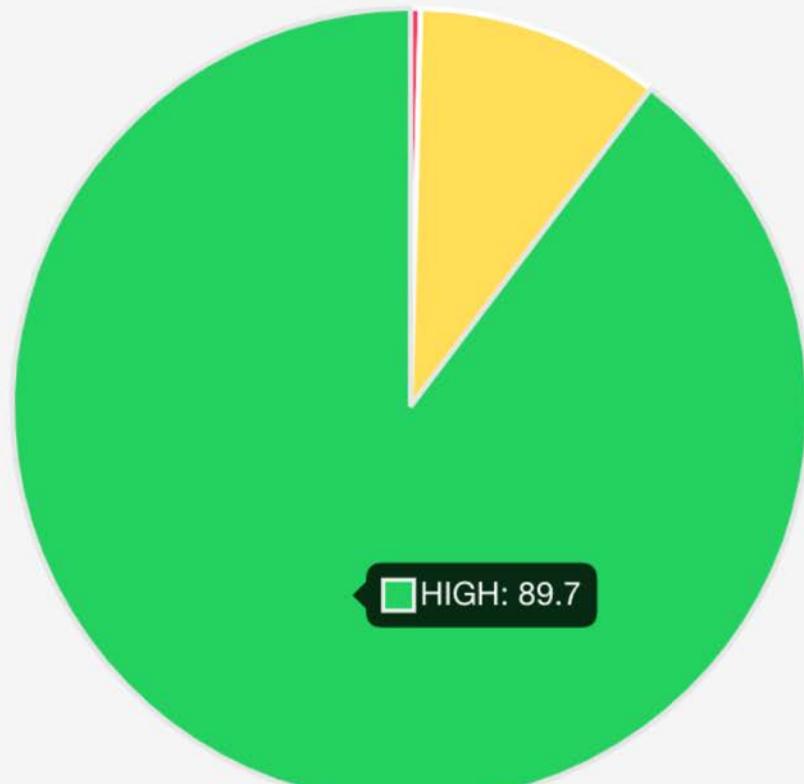
<http://www.tanbih.org>

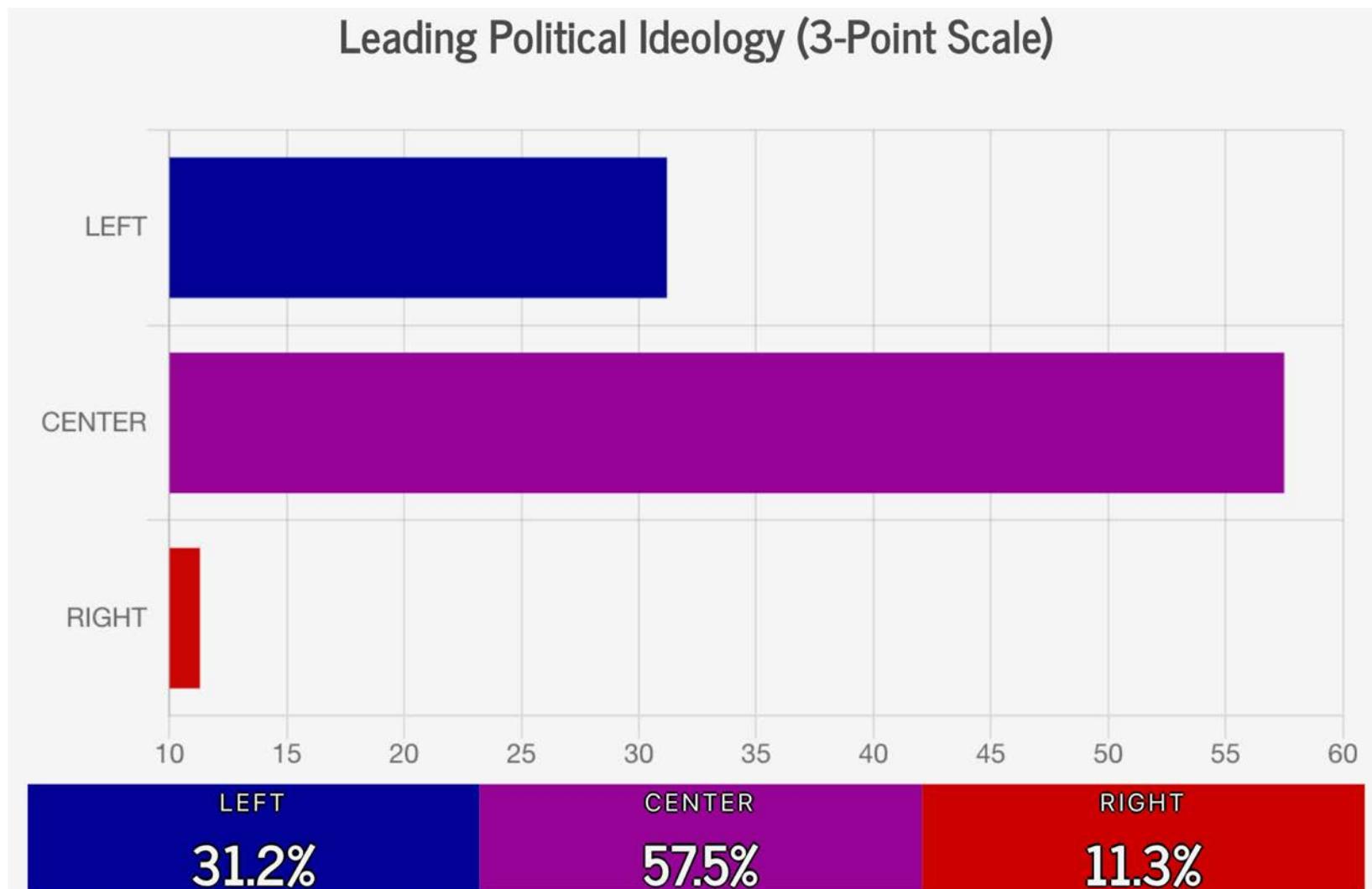


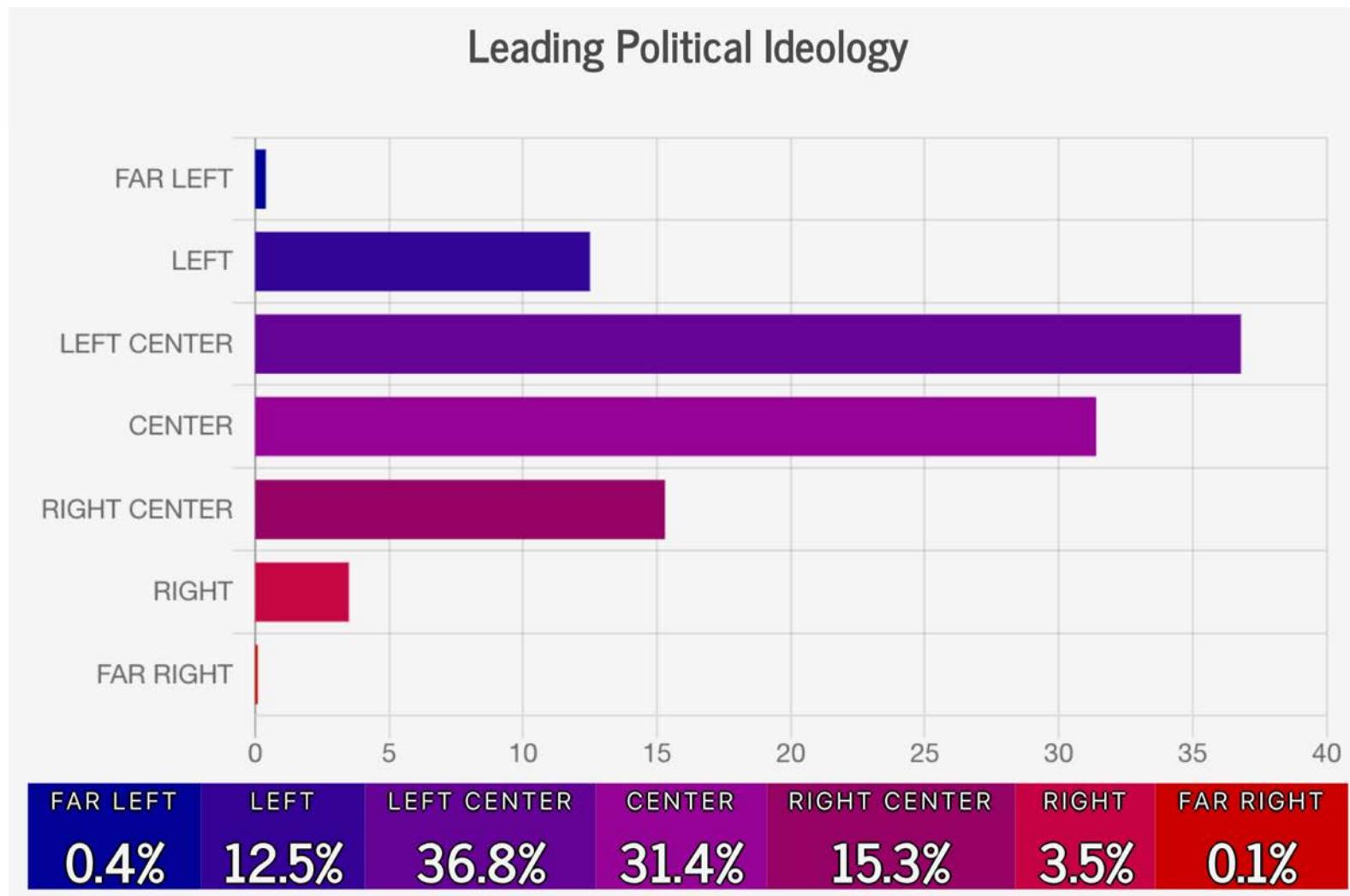
Distribution of General Frame of Reporting



### Factuality of Reporting



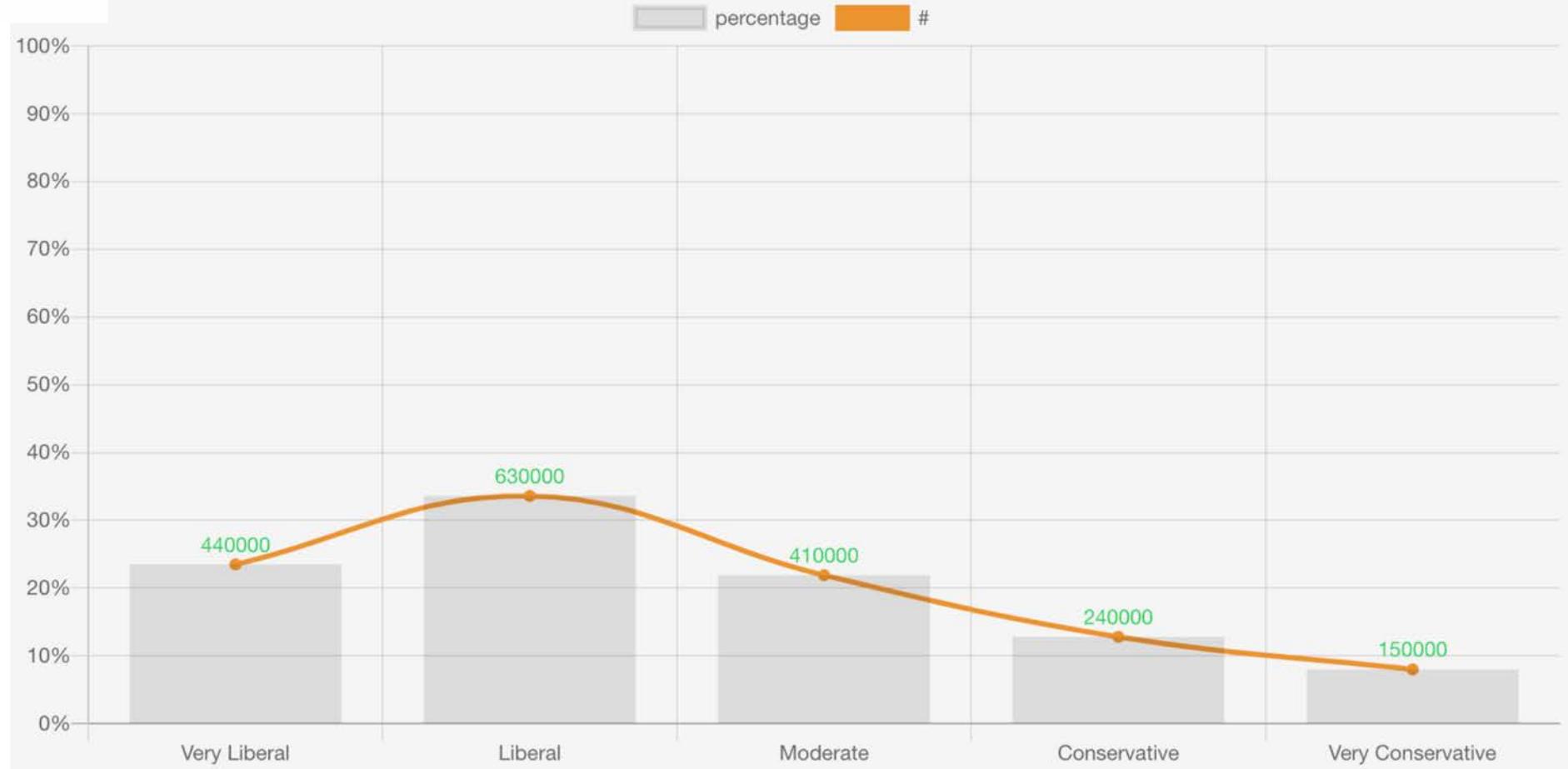




<http://www.tanbih.org>



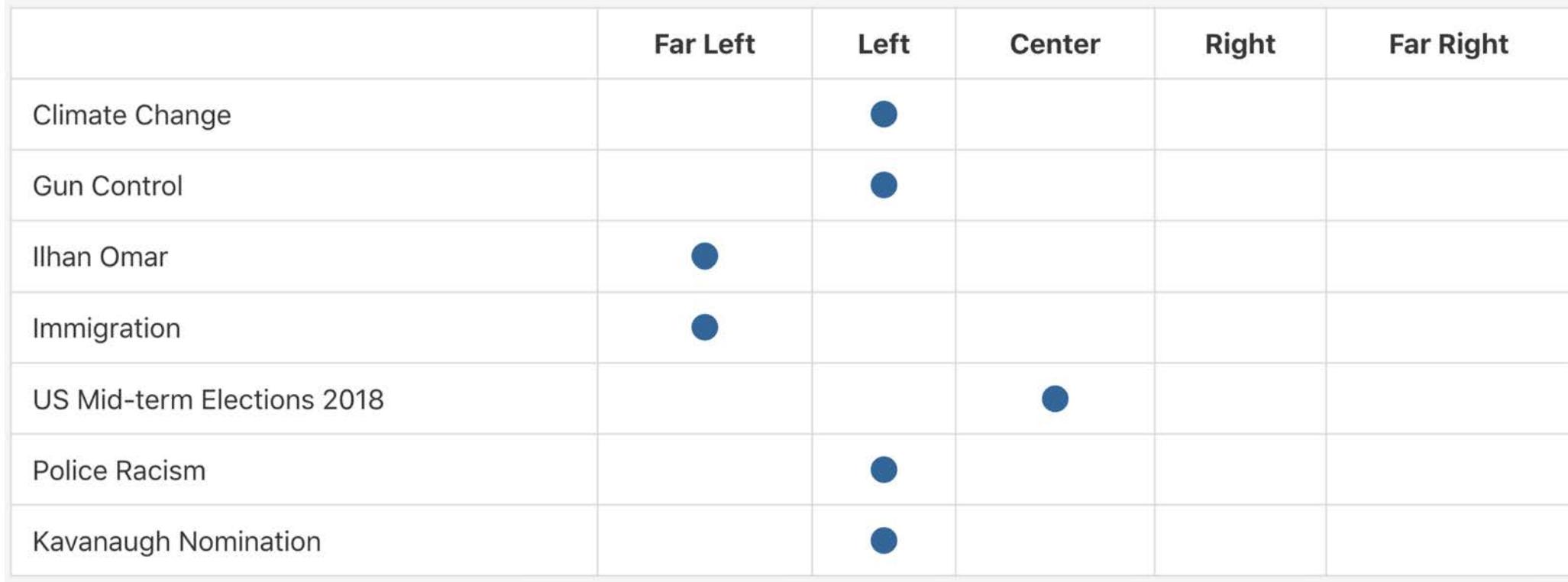
### Audience Reach



<http://www.tanbih.org>



### Twitter User Based Bias Prediction





## Climate Change 69% PRO

### PRO Claims

*The government should increase environmental regulations to prevent climate change (94)*

**78%(74)**

AGREE

**22%(20)**

DISAGREE

### AGAINST Claims

*Climate change is a flat-out hoax (292)*

**3%**

AGREE

**97%(281)**

DISAGREE



Try Tanbih: <http://www.tanbih.org>

*This work is part of the Tanbih project, developed in collaboration between QCRI and MIT-CSAIL, which aims to limit the effect of “fake news”, propaganda, and media bias by making users aware of what they are reading.*

# References (1)

- Atanas Atanasov, Gianmarco De Francisci Morales and Preslav Nakov: Predicting the Role of Political Trolls in Social Media. CoNLL 2019
- Pepa Atanasova, Lluís Màrquez, Alberto Barrón-Cedeño, Tamer Elsayed, Reem Suwaileh, Wajdi Zaghouani, Spas Kyuchukov, Giovanni Da San Martino, Preslav Nakov: Overview of the CLEF-2018 CheckThat! Lab on Automatic Identification and Verification of Political Claims. Task 1: Check-Worthiness. CLEF (Working Notes) 2018
- Pepa Atanasova, Preslav Nakov, Lluís Màrquez, Alberto Barrón-Cedeño, Georgi Karadzhov, Tsvetomila Mihaylova, Mitra Mohtarami, James Glass. Automatic Fact-Checking Using Context and Discourse Information. Journal of Data and Information Quality (JDIQ)
- Pepa Atanasova, Preslav Nakov, Georgi Karadzhov, Mitra Mohtarami, Giovanni Da San Martino: Overview of the CLEF-2019 CheckThat! Lab: Automatic Identification and Verification of Claims. Task 1: Check-Worthiness. CLEF (Working Notes) 2019
- Ramy Baly, Georgi Karadzhov, Abdelrhman Saleh, James R. Glass, Preslav Nakov: Multi-Task Ordinal Regression for Jointly Predicting the Trustworthiness and the Leading Political Ideology of News Media. NAACL-HLT (1) 2019: 2109-2116
- Ramy Baly, Georgi Karadzhov, Dimitar Alexandrov, James R. Glass, Preslav Nakov: Predicting Factuality of Reporting and Bias of News Media Sources. EMNLP 2018: 3528-3539
- Ramy Baly, Mitra Mohtarami, James R. Glass, Lluís Màrquez, Alessandro Moschitti, Preslav Nakov: Integrating Stance Detection and Fact Checking in a Unified Corpus. NAACL-HLT (2) 2018: 21-27
- Alberto Barron-Cedeno, Giovanni Da San Martino, Israa Jaradat, Preslav Nakov. Proppy: Organizing News Coverage on the Basis of Their Propagandistic Content. Information Processing and Management (IPM journal)
- Alberto Barrón-Cedeño, Giovanni Da San Martino, Israa Jaradat, Preslav Nakov: Proppy: A System to Unmask Propaganda in Online News. AAAI 2019: 9847-9848
- Giovanni Da San Martino, Alberto Barrón-Cedeño, Preslav Nakov: Findings of the NLP4IF-2019 Shared Task on Fine-Grained Propaganda Detection. NLP4IF@EMNLP 2019
- Giovanni Da San Martino, Seunghak Yu, Alberto Barrón-Cedeño, Rostislav Petrov, Preslav Nakov: Fine-Grained Analysis of Propaganda in News Articles. EMNLP 2019

# References (2)

- Kareem Darwish, Dimitar Alexandrov, Preslav Nakov, Yelena Mejova: Seminar Users in the Arabic Twitter Sphere. *SocInfo* (1) 2017: 91-108
- Kareem Darwish, Michael Aupetit, Peter Stefanov, Preslav Nakov. Unsupervised User Stance Detection on Twitter. *ICWSM* 2020
- Yoan Dinkov, Ahmed Ali, Ivan Koychev, Preslav Nakov: Predicting the Leading Political Ideology of YouTube Channels Using Acoustic, Textual, and Metadata Information. *INTERSPEECH* 2019
- Yoan Dinkov, Ivan Koychev, Preslav Nakov: Detecting Toxicity in News Articles: Application to Bulgarian. *RANLP* 2019
- Tamer Elsayed, Preslav Nakov, Alberto Barrón-Cedeño, Maram Hasanain, Reem Suwaileh, Giovanni Da San Martino, Pepa Atanasova. Overview of the CLEF-2019 CheckThat!: Automatic Identification and Verification of Claims. *CLEF* 2019
- Pepa Gencheva, Preslav Nakov, Lluís Màrquez, Alberto Barrón-Cedeño, Ivan Koychev: A Context-Aware Approach for Detecting Worth-Checking Claims in Political Debates. *RANLP* 2017: 267-276
- Maram Hasanain, Reem Suwaileh, Tamer Elsayed, Alberto Barrón-Cedeño, Preslav Nakov. Overview of the CLEF-2019 CheckThat! Lab: Automatic Identification and Verification of Claims. Task 2: Evidence and Factuality. *CELF* 2019
- Israa Jaradat, Pepa Gencheva, Alberto Barrón-Cedeño, Lluís Màrquez, Preslav Nakov: ClaimRank: Detecting Check-Worthy Claims in Arabic and English. *NAACL-HLT (Demonstrations)* 2018: 26-30
- Georgi Karadzhov, Preslav Nakov, Lluís Màrquez, Alberto Barrón-Cedeño, Ivan Koychev: Fully Automated Fact Checking Using External Sources. *RANLP* 2017: 344-353
- Daniel Kopev, Ahmed Ali, Ivan Koychev, Preslav Nakov. Detecting Deception in Political Debates Using Acoustic and Textual Features. *ASRU* 2019
- Todor Mihaylov, Georgi Georgiev, Preslav Nakov: Finding Opinion Manipulation Trolls in News Community Forums. *CoNLL* 2015: 310-314
- Todor Mihaylov, Ivan Koychev, Georgi Georgiev, Preslav Nakov: Exposing Paid Opinion Manipulation Trolls. *RANLP* 2015: 443-450
- Todor Mihaylov, Preslav Nakov: Hunting for Troll Comments in News Community Forums. *ACL* (2) 2016
- Todor Mihaylov, Tsvetomila Mihaylova, Preslav Nakov, Lluís Màrquez, Georgi Georgiev, Ivan Koychev: The dark side of news community forums: 84 opinion manipulation trolls. *Internet Research* 28(5): 1292-1312 (2018)

# References (3)

- Tsvetomila Mihaylova, Georgi Karadzhov, Pepa Atanasova, Ramy Baly, Mitra Mohtarami, Preslav Nakov: SemEval-2019 Task 8: Fact Checking in Community Question Answering Forums. SemEval@NAACL-HLT 2019: 860-869
- Tsvetomila Mihaylova, Preslav Nakov, Lluís Màrquez, Alberto Barrón-Cedeño, Mitra Mohtarami, Georgi Karadzhov, James R. Glass: Fact Checking in Community Forums. AAAI 2018: 5309-5316
- Mitra Mohtarami, James Glass, Preslav Nakov. Contrastive Language Adaptation for Cross-Lingual Stance Detection. EMNLP 2019
- Mitra Mohtarami, Ramy Baly, James R. Glass, Preslav Nakov, Lluís Màrquez, Alessandro Moschitti: Automatic Stance Detection Using End-to-End Memory Networks. NAACL-HLT 2018: 767-776
- Preslav Nakov, Tsvetomila Mihaylova, Lluís Màrquez, Yashkumar Shiroya, Ivan Koychev: Do Not Trust the Trolls: Predicting Credibility in Community Question Answering Forums. RANLP 2017: 551-560
- Slavena Vasileva, Pepa Gencheva, Lluís Màrquez, Alberto Barrón-Cedeño, Preslav Nakov: It Takes Nine to Smell a Rat: Neural Multi-Task Learning for Check-Worthiness Prediction. RANLP 2019
- Seunghak Yu, Giovanni Da San Martino, Jim Glass, and Preslav Nakov. Experiments in Detecting Persuasion Techniques in News. NeurIPS 2019 workshop on AI for Social Good.
- Todor Staykovski, Alberto Barrón-Cedeño, Giovanni Da San Martino, Preslav Nakov. Dense vs. Sparse Representations for News Stream Clustering. Text2Story'19@ECIR'19
- Marcos Zampieri, Shervin Malmasi, Preslav Nakov, Sara Rosenthal, Noura Farra, Ritesh Kumar. Predicting the Type and Target of Offensive Posts in Social Media. NAACL-2019
- Marcos Zampieri, Shervin Malmasi, Preslav Nakov, Sara Rosenthal, Noura Farra, Ritesh Kumar. SemEval-2019 Task 6: Identifying and Categorizing Offensive Language in Social Media (OffensEval). SemEval 2019
- Yifan Zhang, Giovanni Da San Martino, Alberto Barrón-Cedeño, Salvatore Romeo, Jisun An, Haewoon Kwak, Todor Staykovski, Israa Jaradat, Georgi Karadzhov, Ramy Baly, Kareem Darwish, James Glass, Preslav Nakov. Tanbih: Get To Know What You Are Reading. EMNLP 2019 (demo)
- Dimitrina Zlatkova, Preslav Nakov, Ivan Koychev: Fact-Checking Meets Fauxtography: Verifying Claims About Images. EMNLP 2019