A Dashboard for Mitigating the COVID-19 Misinfodemic

UNIVERSITY OF TEXAS ARLINGTON

Build and Broaden: Conference on Social Connections to Promote Individual and Community Resilience in Post-COVID-19 Society

Zhengyuan Zhu[1] Kevin Meng* Josue Caraballo[1] Israa Jaradat[1] Xiao Shi[1] Zeyu Zhang[1] Farahnaz Arkami[1] Haojin Liao[1] Fatma Arslan[1] Damian Jimenez[1] Mohammed Samiul Saeef[1] Paras Pathak[1] Chengkai Li[1]

[1]University of Texas at Arlington [2]Massachusetts Institute of Technology

iDiR Innovative Data Intelligence Research Laboratory

System: idir.uta.edu/covid-19

The Quest to Mitigate the COVID-19 Misinfodemic

Social media posts make false and misleading claims

- How the virus is transmitted
- How authorities and people are responding to the pandemic
- COVID-19 symptoms, treatments, and so on

Misinformation exacerbates the vicious impact of the virus

- Drown out credible information
- Interfere with measures to contain the outbreak
- Deplete resources and overloads the health care system

Long-term goals: understand the surveillance of, impact of, and effective interventions against the COVID-19 misinfodemic

- For surveillance, we seek to discover the patterns by which different types of COVID-19 misinformation spread
- To understand the impact of misinformation, we aim to compare the spreading of the SARS-CoV-2 virus and misinformation and derive their correlations

Data Collection: Cases Counts, Tweets, COVID-19 Related Facts

Counts of confirmed cases, deaths, and recoveries

- Johns Hopkins University
- The New York Times (NYT)
- The COVID Tracking Project

A catalog and a taxonomy of COVID-19 related facts.

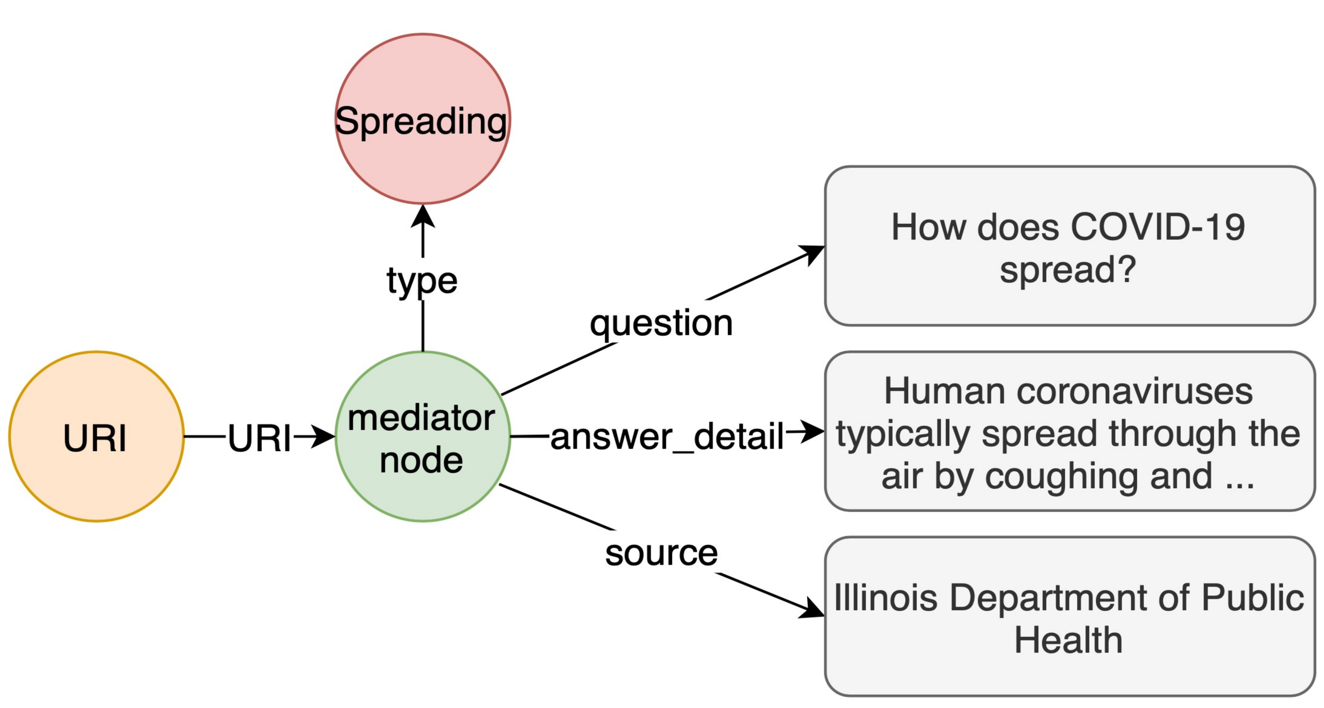
- 9,512 entries from 21 credible website
- Statements from authoritative organizations
- Verdicts, debunks, and explanations of factual claims (of which the truthfulness varies) from fact checking websites
- FAQs both from credible sources (e.g., FDA, NYT)
- All collected data are available at: https://github.com/idirlab/covid19data

Tweets from January 1st to May 16th, 2020

- 250 million COVID-19 related tweets
- Remove tweets that do not have location information, and display the tweets based on locations on the dashboard
- Resulting in 34.6 million remaining tweets

Catalog entries in RDF

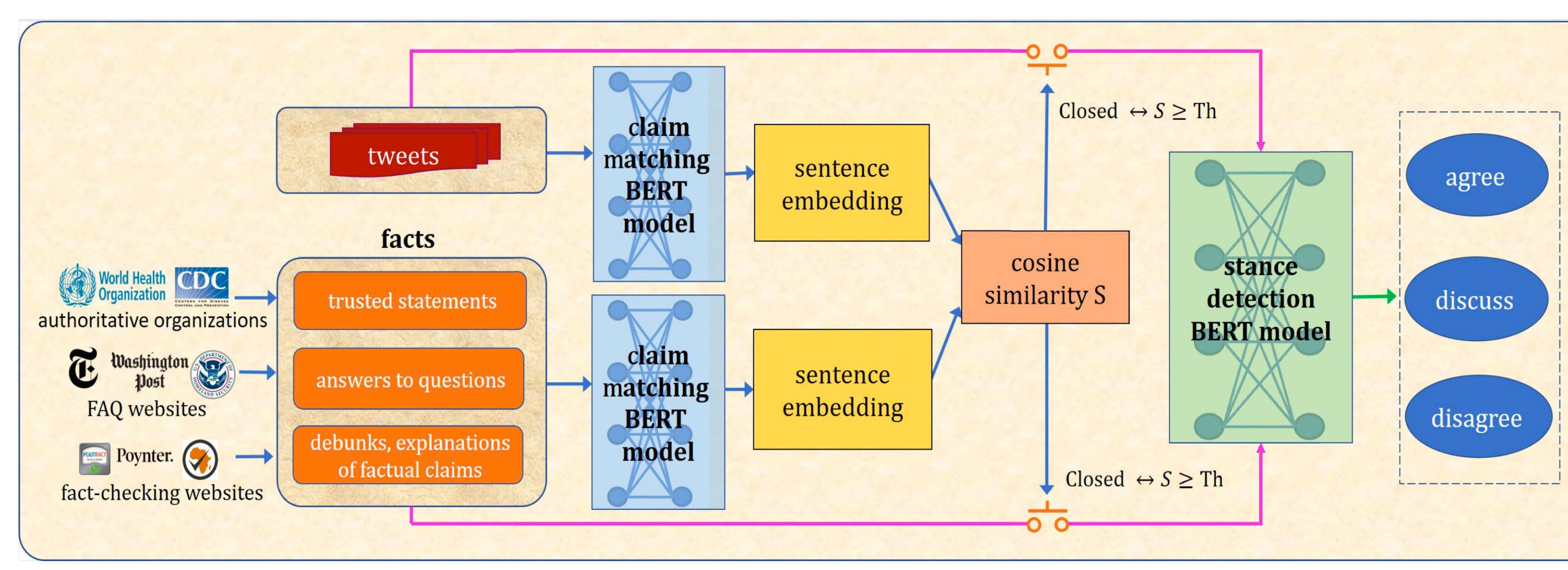
- Dataset: https://cokn.org/deliverables/7-covid19-kg/



Matching Tweets with Facts and Stance Detection

Tweet	Fact	Taxonomy Categories	Similarity	Stance
Coronavirus cannot be passed by dogs or cats but they can test positive.	There has been no evidence that pets such as dogs or cats can spread the coronavirus.	Animals, Spreading	0.817	agree
More people die from the flu in the U.S. in 1 day than have died of the Coronavirus across the world ever.	Right now, it appears that COVID-19, the disease caused by the new coronavirus, causes more cases of severe disease and more deaths than the seasonal flu.	Cases	0.816	disagree

End-to-End Stance Detection System Architecture



Performance of Claim Matching

- Given the catalog of COVID-19 related facts and the tweets, locate a set of tweets that discuss each fact

Metric	@5	@10	@20	@50	@100
Precision	0.80	0.80	0.70	0.56	0.52
nDCG	0.62	0.72	0.78	0.81	0.83

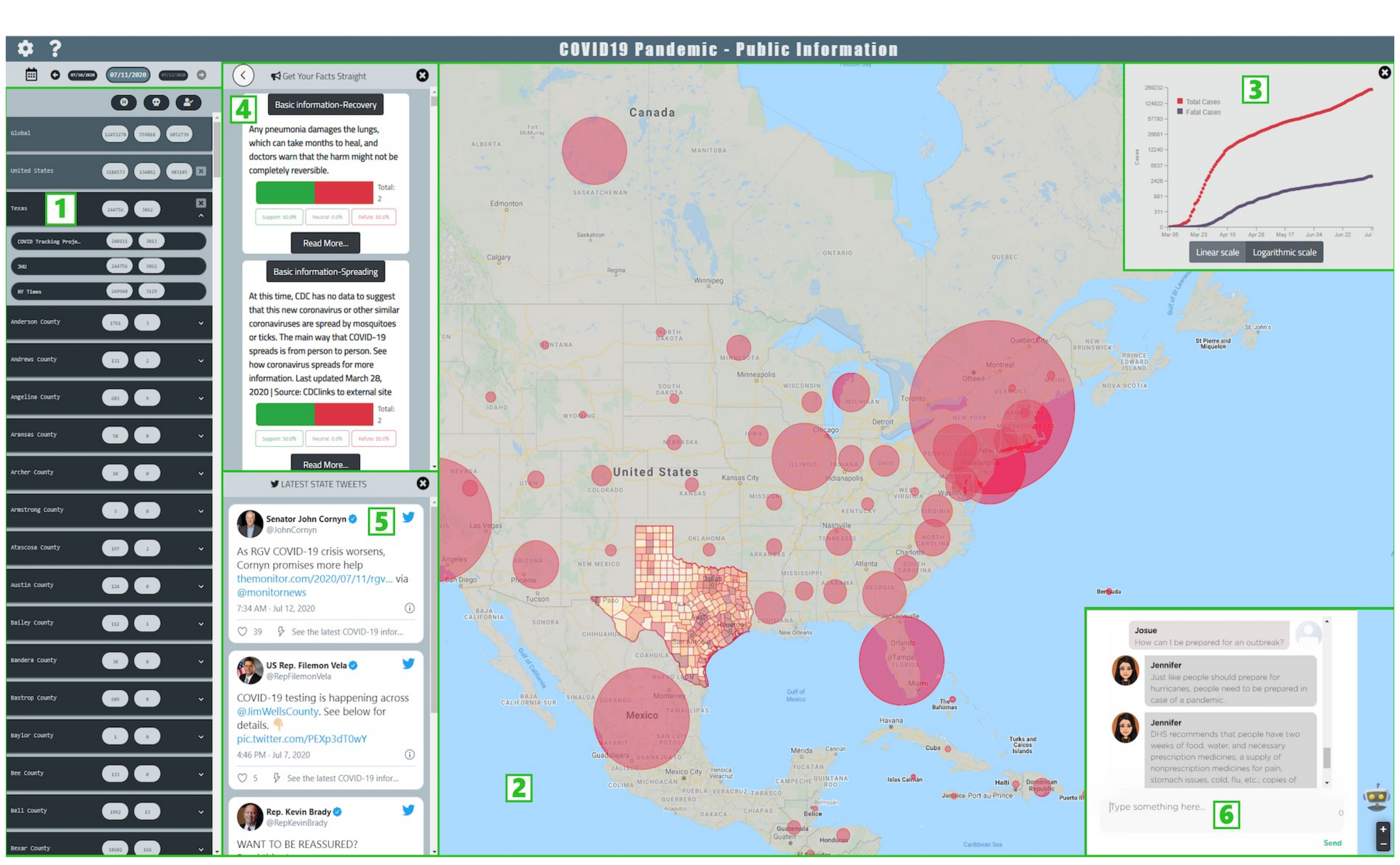
Performance of Stance Detection

- Detect tweets' stance toward fact

Model	F1 score			
	agree	discuss	disagree	macro
Stance-BERT$_{window}$(FNC-1)	0.65	**0.45**	**0.84**	**0.65**
Stance-BERT$_{trunc}$(FNC-1)	**0.66**	0.41	0.82	0.63
(Xu et al., 2018)(FNC-1)	0.55	0.15	0.73	0.48
Stance-BERT$_{window}$(COVID-19)	0.75	0.03	0.58	0.45

COVID-19 Dashboard GUI

Components: 1) Geographic region selection panel. 2) Interactive map. 3) Timeline chart. 4) Panel of facts. 5) Government tweets. 6) Chatbot.



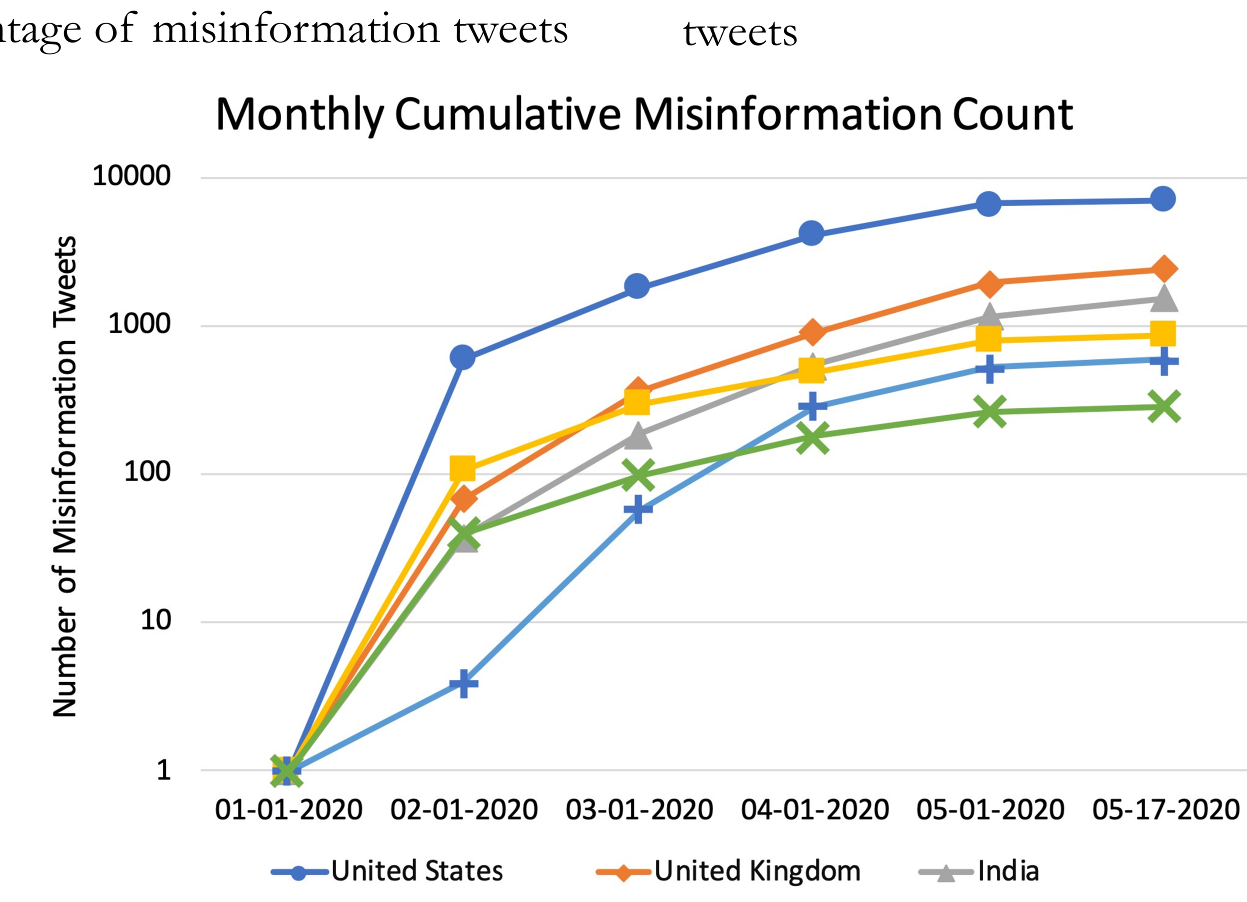
Case Study: COVID-19 Misinformation Analysis

Country	Confirm	Death	Recover
United States	0.763	0.738	0.712
United Kingdom	0.862	0.833	-
India	0.794	0.798	0.755
Canada	0.706	0.667	0.663
Australia	0.954	0.922	0.887
Philippines	0.720	0.696	0.618

Correlation between the percentage of cases and the percentage of misinformation tweets

Category	Count	Percentage
Definition	2503	15.1
Spreading	2118	12.7
Other	1450	8.7
Testing	1301	7.8
Disease Alongside	936	5.6
Total	8308	49.9

Most frequent categories of misinformation tweets



* The work was completed while Kevin Meng was a UTA affiliate.