VACCINELies : A Natural Language Resource for Learning to Recognize Misinformation about the COVID-19 and the Human Papillomavirus Vaccines

Maxwell A. Weinzierl and **Sanda M. Harabagiu**
Human Language Technology Research Institute
University of Texas at Dallas

The Problem:

Misinformation spreading, especially on social media, is believed to be responsible for vaccine hesitancy (Kouzy et al., 2020). It is imperative for public health practitioners to know *what misinformation is spreading* as well as *who is adopting or rejecting it*, such that interventions can be tailored appropriately. Public health messaging approaches could not only inoculate against misinformation, but also effectively reach social media users with promise to shift or bolster vaccine attitudes.

❑Current language resources (e.g. PHEME, Twitter 15, Twitter 16) do not cover vaccine misinformation. We were inspired by COVIDLies (Hossain et al 2020), which annotated tweets evoking 86 known misconceptions about COVID-19, and created VACCINELies, an annotated dataset addressing misinformation about:

➢ *The Covid-19 vaccines*
➢ *The Vaccines against Human Papillomavirus (HPV)*

Our Solution:

VACCINELies consists of:
◁1▷ Misinformation Targets (**MisTs**) , addressing misinformation towards COVID-19 or HPV vaccines;
◁2▷ the **tweet IDs** for those tweets that were judged as evoking any of the MisTs available in VACCINELies;
◁3▷ annotation of the **stance** of each tweet author that evoked a MisT, indicating if they *Accept* the Mist; Reject it, or they have no stance towards it.
◁4▷ a **taxonomy of the MisTs**, which enables the interpretation of the themes and concerns characterizing the vaccine misinformation available in **VACCINELies**.
VACCINELies can be considered as consisting of two datasets: **COVAXLies** and **HPVAXLies**

Statistics in VACCINELies

	CoVaxLies	HpVaxLies	Total
MisTs	48	21	69
Evoke	7,152	2,230	9,382
Accept	3,720	1,365	5,085
Reject	2,194	617	2,811
No Stance	1,238	248	1,486
Tweets	12,118	2,524	14,642

(A) CoVaxLies
(B) HpVaxLies
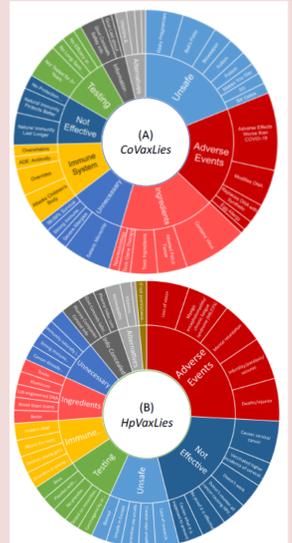
Creating VACCINELies :

◁*STEP 1*▷ Discovering Misinformation Targets (**MisTs**)
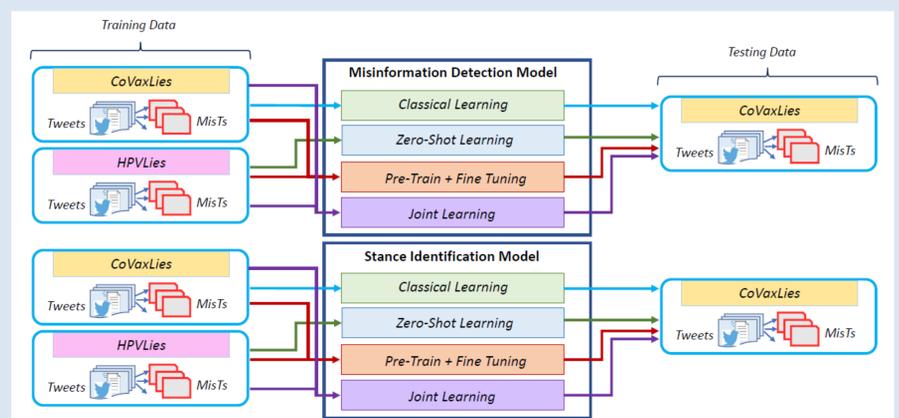◁*STEP 2*▷ Generating a Taxonomy of Vaccine Misinformation
◁*STEP 3*▷ Identifying Tweets evoking MisTs Using BM25 and a BERT-RERANK system
◁*STEP 4*▷ Annotating the *Stance* of tweets evoking each MisT

Transfer Learning:

The Results:
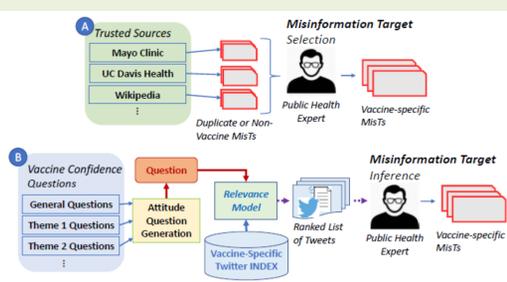
Vaccine misinformation detection results

Results of the BERT Vaccine Misinformation Evocation Detector (BERT-VMED) utilizing vaccine transfer learning scenarios:

	CoVaxLies	HpVaxLies	Total
MisTs	48	21	69
Evoke	7,152	2,230	9,382
Accept	3,720	1,365	5,085
Reject	2,194	617	2,811
No Stance	1,238	248	1,486
Tweets	12,118	2,524	14,642

Stance Identification results

Results of the BERT Vaccine Misinformation Stance Identifier (BERT-VMSI) utilizing vaccine transfer learning scenarios:

Testing	Scenario	Macro F_1	P	R	Accept F_1	P	R	Reject F_1	P	R
CoVaxLies	Classical	83.4	81.6	**85.5**	85.9	81.5	**90.8**	80.9	81.8	**80.1**
	Zero-Shot	72.5	84.1	64.2	79.0	85.4	73.5	65.9	82.9	54.8
	Joint	83.3	82.8	84.1	85.2	81.9	88.8	**81.4**	**83.6**	79.4
	Pre-Train	**83.6**	**85.7**	81.5	**86.8**	**88.7**	85.0	80.3	82.7	78.1
HpVaxLies	Classical	79.6	79.9	79.4	83.1	82.2	84.0	76.2	77.7	74.8
	Zero-Shot	74.6	71.8	79.7	79.2	68.4	**93.9**	70.0	75.3	65.4
	Joint	80.5	80.5	80.6	85.9	83.8	88.2	75.0	77.2	72.9
	Pre-Train	**84.0**	**85.1**	**83.2**	**88.1**	**86.4**	89.7	**80.0**	**83.7**	**76.6**