



Analysing the Spread of Fake News:
**Raising Awareness &
Educating the Public with a
Simulation Tool**

GT17

Lee Cheng Leng

Tho Shu Ting Belinda

Lim Zi Yuan

Sean Kwek Si Wei

Joel-David Wong



Real or Fake?

STATEMENT

“University offers students a grave in the soil to lie in to cope with exam stress”



Real ~~or Fake?~~

STATEMENT

REAL

“University offers students a grave to lie in to cope with exam stress”

University offers students grave to lie in to cope with exam stress

Radboud University in Nijmegen has dug a 'purification grave' to help students reflect on the futility of their existence - and it's certainly different

SHARE [f](#) [t](#) [in](#) [r](#) [e](#) [c](#) COMMENTS

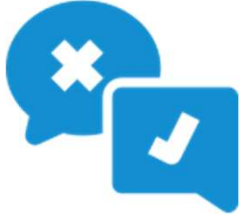
By **Hannah Dodd** Video Journalist
14:55, 9 NOV 2019

NEWS



DUTCH UNIVERSITY LETS STUDENTS LIE IN A GRAVE ON CAMPUS TO DE-STRESS





Real or Fake?

STATEMENT

“Woman arrested for defecating on boss’ desk after winning the lottery”



Real or Fake?

STATEMENT

“Woman arrested for defecating on boss’ desk after winning the lottery”

FAKE

How
defec
\$3m
going

- Dave We
- He said h
- A bogus lottery b



Jesus Chrysler
@JesusChrysler15

Woman arrested for defecating on boss’ desk after winning the lottery

thevalleyreport.com/2016/04/25/wom... via @theval

<https://thevalleyreport.com/2016/04/25/woman-arrested-for-defecating-on-boss-desk-after-winning-the-lottery/>



Woman arrested for defecating on boss’ desk after winning the lottery
DAVE WEASEL NEW YORK – A 41-year-old woman had the winning lottery ticket worth over 3 million dollars on Friday night, but showed up to work ...
thevalleyreport.com

7:15 AM · Aug 31, 2019 · Twitter Web App

106 Retweets 422 Likes



Real or Fake?

STATEMENT

**“The word ‘Racism’ is
searched more often than
‘Thesaurus’ ”**



~~Real or~~ Fake?

STATEMENT

FAKE





Real or Fake?

STATEMENT

“Woman from China opens plane emergency door, says it was too stuffy”



Real ~~or Fake?~~

STATEMENT

REAL

“Woman from China opens plane emergency door, says it was too stuffy”

Woman from China opens plane emergency door, says it was too stuffy

Probably shouldn't have done that.

Syahindah Ishak | September 25, 06:14 pm

f 6.6K





Background & Motivation

- 4 in 5 Singaporeans are confident in spotting fake news but 90% are actually wrong when they are put to the test
- Effects of fake news can lead to increased polarization and potential conflicts among the public
- Protection from Online Falsehoods and Manipulation Act (POFMA) was enacted in 2019 to prevent electronic communication of falsehoods



Objectives



1) Study the various methods employed by spreaders of fake news by diving into the topic of vaccination



2) Educate users of the consequences of spreading fake news



3) Develop an interactive tool which can be used as an education tool on fake news by allowing users to experience the following:

- Techniques employed by spreaders of fake news
- How quickly fake news can spread
- Consequences of spreading fake news

CLIENT



Singapore Press Holdings

Initiatives:

1. **FAKE NEWS DEBUNKED**

Manually check the validity of questionable sources

2. **Straits Times Media Analysis Course**

→ Improve critical analysis of news sources and their credibility

THE STRAITS TIMES



SINGAPORE

POLITICS

ASIA

WORLD

VIDEOS

MULTIMEDIA

LIFESTYLE

FOOD

FORUM

OPINION

FAKE NEWS DEBUNKED



Video of mass brawl involving workers did not happen in S'pore, say police

Oct 1, 2019, 5:07 pm



HOME

ABOUT US

ENGLISH

DIGITAL

MEDIA

ARTICLES

CONTACT US

THE STRAITS TIMES MEDIA ANALYSIS COURSE

DESCRIPTION

TRAINER

Be equipped with the skills to spot fake news and analyse arguments in the media

Who should attend

- Professionals who deal with or disseminate information at work
- Those who wish to use social media for their benefit

In this course, participants will learn

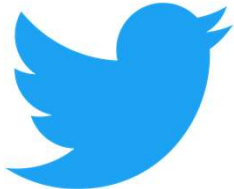

Fighting Fake News

- Impact of the new fake news law

COURSE FEATURES

PROVIDER:	The Straits Times
VENUE:	Level 8 Tower Block, SPH News Centre 1000 Toa Payoh North
DURATION:	10am to 6pm
UPCOMING DATES:	Not Available

DATASET

	Focused Topic:	<u>Analysis on the topic of Vaccination</u>
Tweets timeline	2018 ————— 2019	
Method	 Twitter Scrapping	#vaccines
Results	140,649 unique tweets	



Assumptions & Limitations

Focus of analysis on vaccination

- Might not be able to be generalised to other topics where fake news is rampant

Twitter's suspension of accounts found to be spreading fake news

- Could not study the most influential tweeters in the network as their accounts were suspended

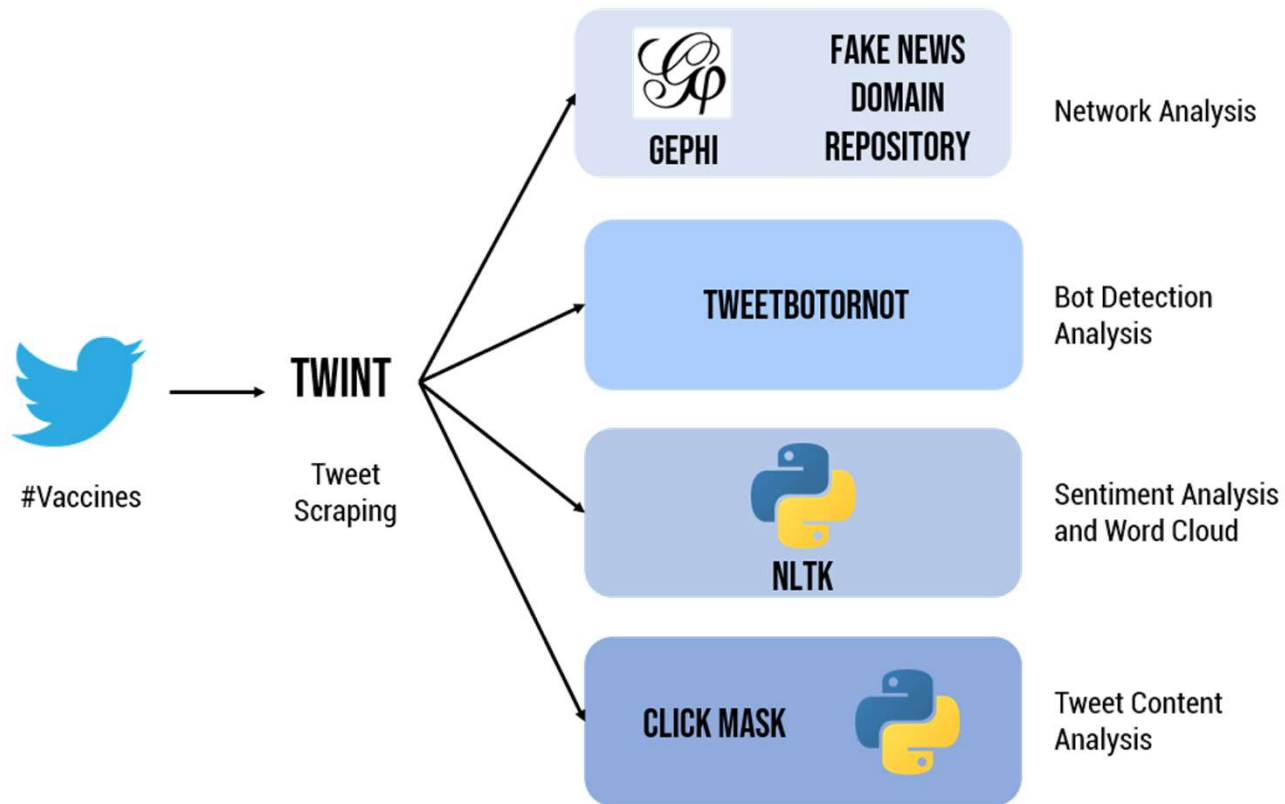


Tools & Technologies

Twint	<ul style="list-style-type: none">•Retrieve tweets for analysis
Gephi	<ul style="list-style-type: none">•To identify the most influential yellow journalists•Study social dynamics in the network
Natural Language Toolkit (nltk)	<ul style="list-style-type: none">•Analyze sentiments of the twitter comments•Generate word clouds
Twitter API	<ul style="list-style-type: none">•A tool to complement the Bot Detection Model
Fake News Domain Repository	<ul style="list-style-type: none">•Retrieve a blacklist of fake news websites
AWS EC2 Server	<ul style="list-style-type: none">•Host Twitter crawler
TweetBotorNot	<ul style="list-style-type: none">•Bot detection classifier model
ClickMask	<ul style="list-style-type: none">•Clickbait factor detection



Architecture Diagram

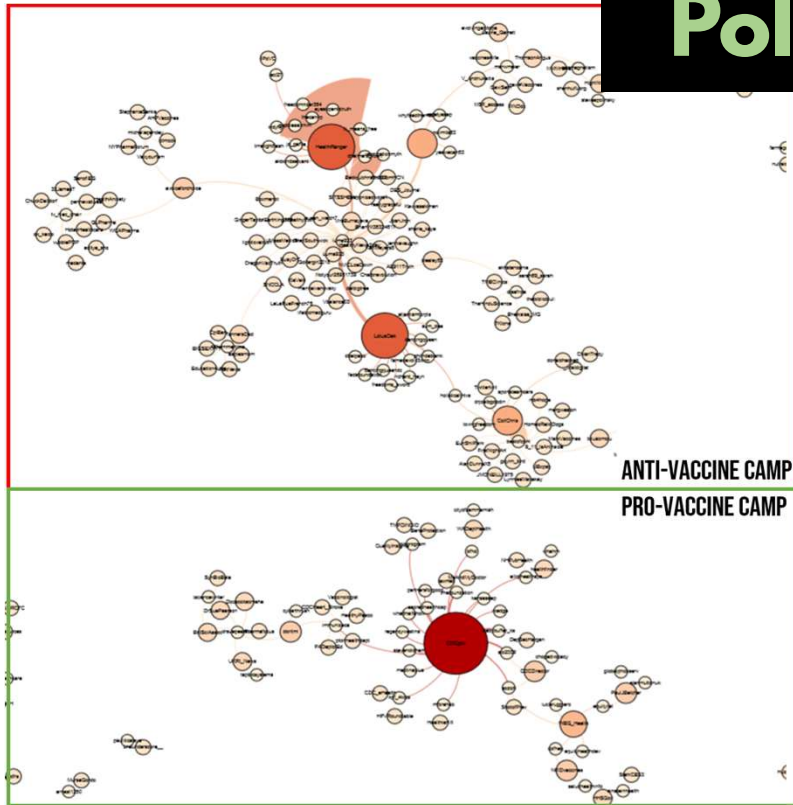


01. **Network Analysis**

- Utilized Gephi to visualize the social dynamics in the spread of fake news
- Conducted retweet analysis to model the spread of information within the #vaccines hashtag
- Identified influential yellow journalists such that we can study their techniques utilized in spreading misinformation

01. Network Analysis

Polarisation of Communities

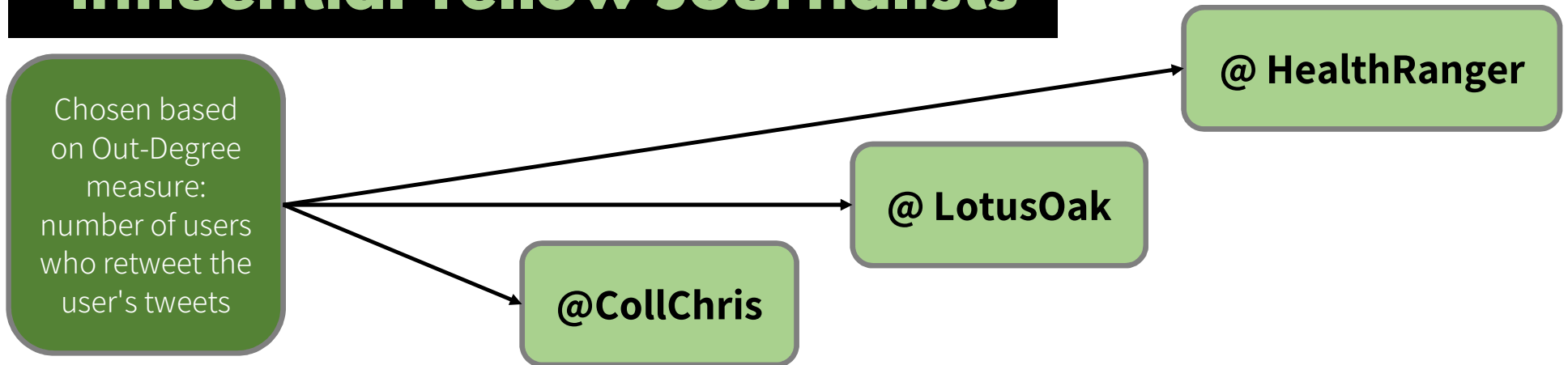


**Pro-Vaccine and Anti-Vaccine
Tweeters are disparate
communities**

**"Filter Bubble Effect":
ideological polarisation due
to content personalisation
algorithms**

01. Network Analysis

Influential Yellow Journalists



Account suspended

This account has been suspended. [Learn more](#) about why Twitter suspends accounts, or [return](#) to your timeline.



LotusOak @LotusOak2 · 2m

Babies Die in [#India](#) After Getting 5-in-1 Vaccine

thevaccinereaction.org/2019/04/babies...

Two more babies, aged 6 & 8 weeks, died in India following pentavalent [#vaccination](#). Another 32 children came down w/ fever, at least 4 of them were in serious condition.

[#LearnTheRisk](#) [#VaccineDeath](#)



@LotusOak

@CollChris



Chris Collins @CollChris · Oct 30

Vaccines are the car accident. [#edu](#) [#education](#) [#HPVvaccine](#) [#hvp](#) [#flu](#) [#flushot](#) [#mmr](#) [#vaxxed](#) [#vaccines NO](#) [#vaccineswork](#)

Jon Hislop, MD, PhD @HislopMD

People: "I haven't gotten the flu shot in 30 years & I never had the flu!!!"

Okay...but the next time I drive my car I'm putting my seatbelt on.....



01. Network Analysis

Susceptible Tweeters

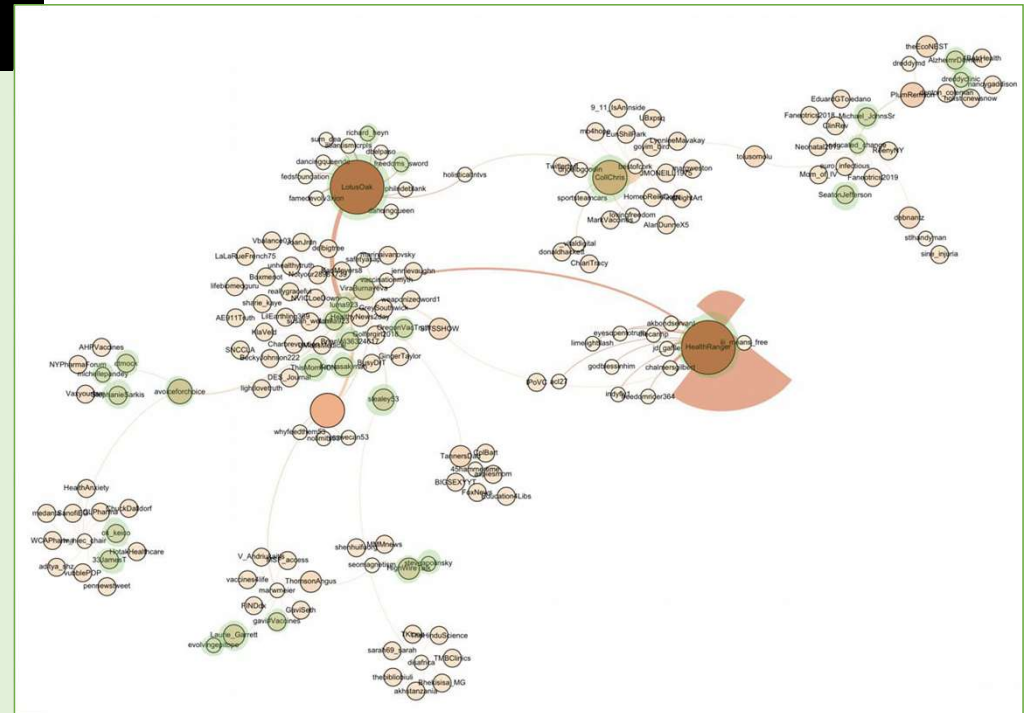
Aim: to identify the number of users who fall prey to misleading claims

- Utilized a fake news domain repository to distinguish yellow journalists from the other tweeters in the network
- Users who retweet the tweets containing fake news but are not originators of these tweets are seen as "susceptible"
- They are target users of our game as these naïve users can actually be charged under the POFMA law

01. Network Analysis

Susceptible Tweeters

79.4% of the anti-vaccine camp are identified as "susceptible"



02. Bot Detection

Bots: Artificial users which simulate the behavior of human users to spread fake news

Important role in spreading of fake news

- Making the content appear more popular using amplification bots



Used a model that takes in bio, location, number of followers and friends to determine if a user is bot.

Since the model has 91.78% accuracy, we determined that **the user is a bot if the model's resultant probability is > 60%.**

02. Bot Detection

35	SalomnM	314024921	0.644
36	rusticrobin	280705761	0.658
37	Just_Truth_Now	1058537186239365120	0.658
38	Lacernman	224111514	0.669
39	soybean61	205838596	0.693
40	MurphyLaw23	700293308	0.707
41	bentonlee1	1409535590	0.708
42	PXPSecurityInve	3305508632	0.742
43	Scrappy94546226	1147658480297369602	0.745
44	1n2by2023	718989462978453504	0.747
45	MelissaShaffe18	2956423968	0.769
46	brazencoal	978907383266267136	0.778
47	Lapidostrous	860152754764791808	0.794
48	piersar62	710579736519512064	0.796
49	kevin9359	913225194	0.803
50	cjsienna55	317223956	0.813
51	LotusOak2	424664120	0.818
52	BridgetAshmore	377474863	0.826
53	RonSowers1	822870565866258433	0.828
54	ellen78834490	948055404700930048	0.847
55	butler_613	784376029892739076	0.850

56	SNCCLA	116289807	0.856
57	DeZagarra	1089296994038886400	0.874
58	Kateri60270481	1144105477216886784	0.881
59	JONJONLIVES1	1069994411805851653	0.886
60	nxumalo_terence	845599446259961856	0.904
61	marymorton9	552331554	0.905
62	susan_welch7	1472540850	0.913
63	PL14	395491775	0.915
64	laughwhenucry	778427264916566016	0.917
65	sharpfang	22887776	0.937
66	porridgeisgood	2800466323	0.941
67	svagdis	2440620382	0.953
68	eclectelectric	2192010002	0.956
69	Paquita_337	886291360692752386	0.957
70	QLover18	947144115753545728	0.962
71	Team_Mary_	2747789530	0.966
72	People1stPlanet	35813556	0.971
73	Freewheeln	32292018	0.974
74	aspiesmom	341678317	0.981
75	swhotmess	3001664273	0.981

@LotusOak

**7,435 out of 14,805
were bot followers**

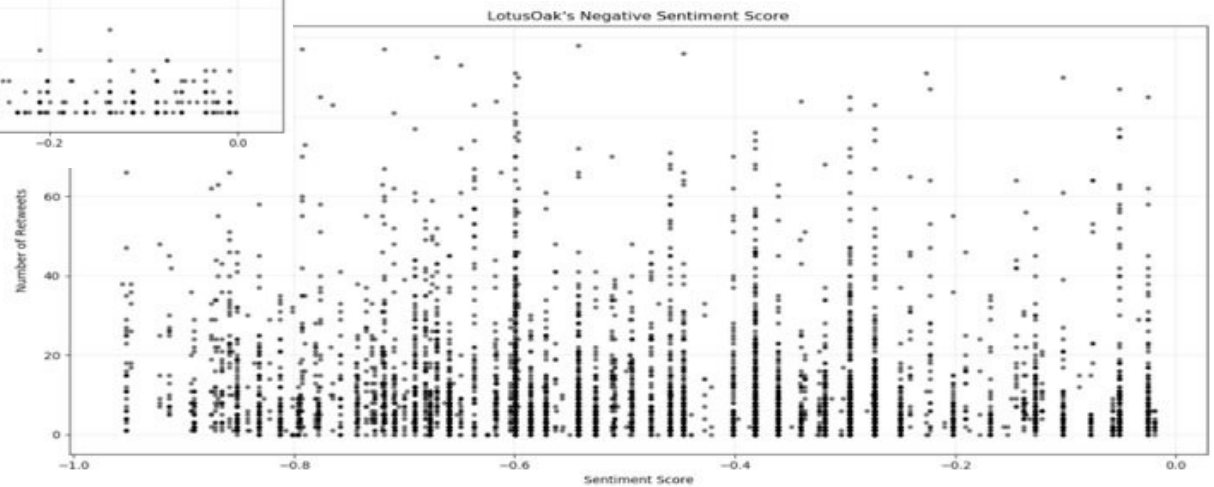
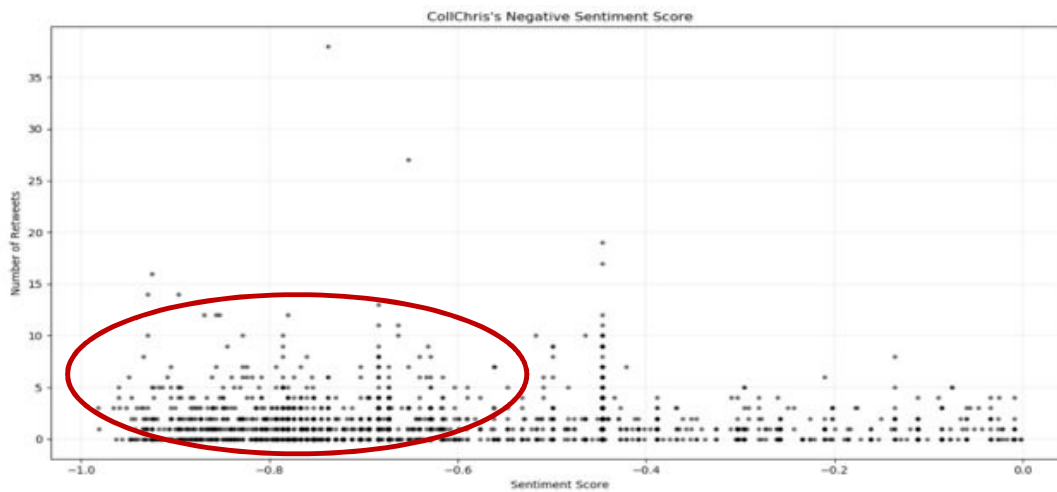
**Tweets with 75
retweets had 41
retweeters who are
bots**

03. Sentiment Analysis

Influencer	Overall Sentiment Score	Strong Sentiment (-0.5 to -0.9999999)		Weak Sentiment (-0.00001 to -0.4999)	
		Tweet Amount	Average Retweet	Tweet Amount	Average Retweet
CollChris	-0.4165	659	0.4709	1019	0.8638
LotusOak2	-0.1015	2554	2.384	2737	2.8385

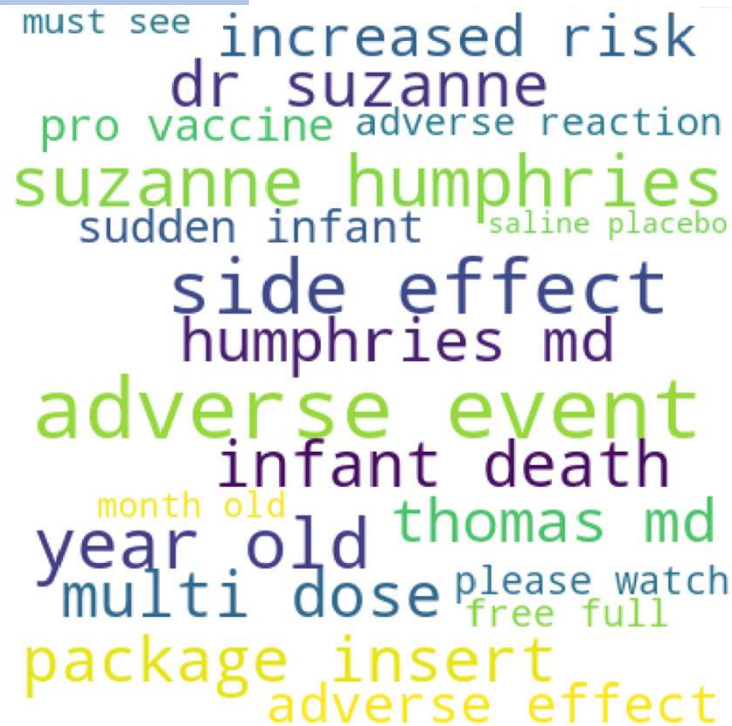
- In General, yellow journalists are **more inclined towards tweeting with negatively worded phrases.**
- However there is **no direct relationship** between sentiment score and retweet count.

03. Sentiment Analysis



03. Word Cloud

LotusOak



A word cloud for LotusOak featuring various terms related to medical research and vaccine safety. The words are arranged in a dense, overlapping cluster. The most prominent words are 'suzanne humphries', 'adverse event', 'side effect', 'infant death', 'multi dose', 'package insert', 'adverse effect', 'year old', 'month old', 'sudden infant', 'saline placebo', 'must see', 'increased risk', 'dr suzanne', 'pro vaccine', 'adverse reaction', 'humphries md', 'thomas md', 'please watch', 'free full', and 'cancer'.

must see increased risk
dr suzanne
pro vaccine adverse reaction
suzanne humphries
sudden infant saline placebo
side effect
humphries md
adverse event
infant death
month old year old thomas md
multi dose please watch
package insert free full
adverse effect

CollChris



A word cloud for CollChris featuring various terms related to vaccine safety and public health. The words are arranged in a dense, overlapping cluster. The most prominent words are 'kid', 'pharma', 'don't', 'health', 'death', 'risk', 'cause', 'say', 'ill', 'cant hear know', 'autism', 'boos', 'sellout', 'cancer', 'child', 'us', 'year old', 'month old', 'sudden infant', 'saline placebo', 'must see', 'increased risk', 'dr suzanne', 'pro vaccine', 'adverse reaction', 'humphries md', 'thomas md', 'please watch', 'free full', and 'cancer'.

kid ill cant hear know
autism cause death
risk say
pharma health don't child
boos sellout cancer
year old month old sudden infant saline placebo
must see increased risk
dr suzanne
pro vaccine adverse reaction
humphries md
thomas md
please watch
free full
cancer

03. Word Cloud

Unique Observations

- Analyzed common themes between tweets of influential tweeters
- **Fear mongering**: commonly used words include death, autism and infant death
 - Causes the reader to focus more on the misinformation presented, rather than their preconceived stance towards vaccines by causing anxiety.
- Words that **require domain knowledge** are present
 - Suzanne Humphries (anti vax doctor), pharma (big pharma controversy) and saline placebo (using saline placebo instead of vaccines supposedly show no changes)
 - Used as a means to support their claims

04. Content Analysis

Clickbait Factor

- Pre-trained model named Click Mask with a model accuracy of 82%
- Run the model to return the probability of a text containing clickbait content
- Compared the clickbait factor between top 3 retweeted tweets and bottom 3 retweeted tweets for LotusOak

04. Content Analysis

Clickbait Factor

Retweet count	Tweet	Clickbait Factor
1630	Think your #vaccines are safe? GREAT! Now let's repeal the 1986 Act that removed liability for vaccine injuries and deaths! Make vaccine manufacturers liable again for the harm caused by their products, only seems fair and just.	<div>It's Clickbait.</div> <div>We are 98.98% sure that this is clickbait.</div>
3	Randomized placebo-controlled #STUDY on adverse effects of #Rubella immunization in seronegative women	<div>Not Clickbait.</div> <div>We are 85.73% sure that this is not clickbait.</div>

Tweets with higher retweet count tend to be flagged as "clickbait" by the model

04. Content Analysis

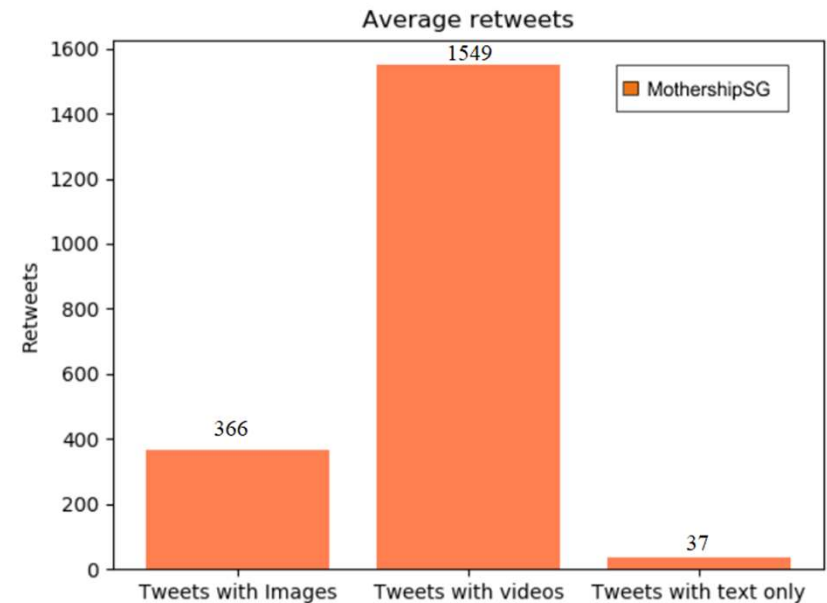
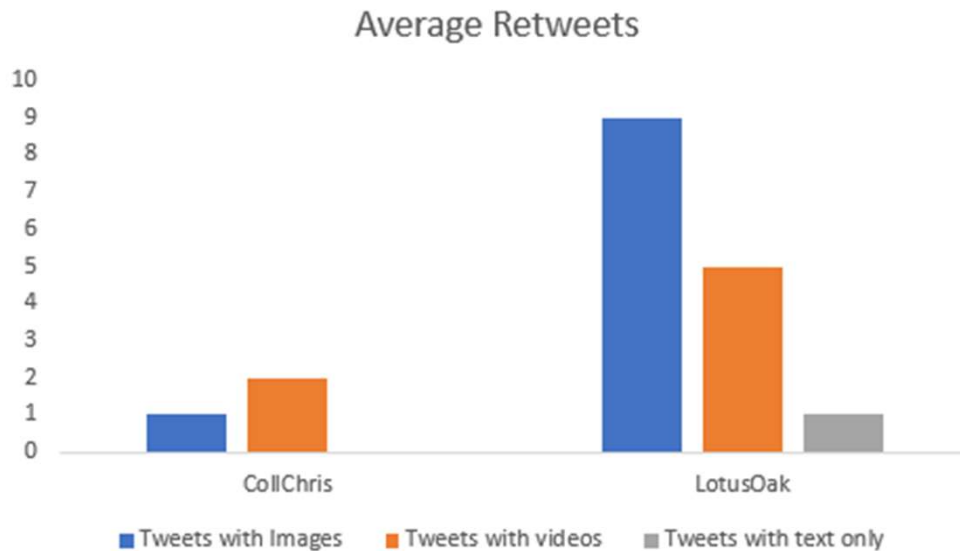
Clickbait Factor

- Yellow journalists tend to focus on "clickbait" titles and emotionally-charged words to incite immediate action
- 140-character limit for tweets means tweeters would not focus on providing factual arguments
- Focus instead on making their tweets believable and sensational to increase its popularity

04. Content Analysis

Content Type

Our findings are consistent with a Twitter study conducted in 2014, where it was found that photos and videos result in higher engagement of 35% and 28% respectively.



TWITTER'S EFFORTS

Suspension of Accounts

- Twitter started banning accounts of fake news spreaders
- Yellow journalists create new accounts under similar names to spread their agenda
- For example, LotusOak was found to have at least 2 other accounts, most of which are suspended
ViraBurnayeva, LotusOak2

TWITTER'S EFFORTS

Fact-Focused Strategy

Know the facts

To make sure you get the best information on vaccinations, resources are available from the Ministry of Health and the Health Promotion Board.

[Visit National Immunisation Registry](#)

[Visit Health Hub](#)

- Shows links of credible sources when users search for a particular topic
- This approach would only be useful if the user is searching up information
- A user who encounters fake news on their timeline might not notice that it is not based off facts

Fake The Noose



- Interactive game to increase awareness of how fake news is spread
- Using gamification to bring about an engaging and educational experience

Fake The Noose



fakethenoose.com

The screenshot displays the fakethenoose.com interface. At the top, a teal button says "I will keep that in mind". Below it, a text box prompts: "Let's move onwards to strengthen the impact by exploring how choices of words can affect readers." Another teal button says "Aye!". Below that, a text box asks: "Take a look at this tweet. Are there any ways we can change the choice of words to make it even more sensationalized?". The tweet is from LotusOak (@LotusOak) and reads: "The government is robbing citizens of their livelihoods! Give us back our e-scooters! #stoptheban #hate #useless". Below the tweet, a text box says "Your newly improved tweet has". To the right, a white box shows "Retweets 221". Below that, a section titled "Real World Effect" explains: "Automated Twitter accounts, called bots, helped spread bogus articles during and after the 2016 U.S. presidential election by making the content appear popular enough that human users would trust it and share it more widely." At the bottom right, a "Badges" section shows five icons: a blue shield with a white 'i', a yellow shield with a sad face, a yellow shield with a play button, a yellow shield with a robot head, and a grey shield with a star.

Users take on the role of a yellow journalist

- Main objective: Spread fake news to the widest audience possible
- First-hand experience of the thought process and techniques used

Instant feedback is provided for each action chosen

- Aids the user in understanding the negative repercussions of their actions and their impact

Inclusion of badges for milestones in the user's fake news journey

- Drive strong behavioral change
- Educate users to be wary of what they read online

Thank You

Any questions?
