

“Longform” Twitter:  
How Twitter threads are used to share  
COVID-19 information

Katy Ilonka Gero, Jennifer Lee, Sarah Huang, Lydia B. Chilton  
Department of Computer Science, Columbia University

CUEPID Data Challenge, April 2020

# How does “longform” info spread on Twitter?

01 Twitter is an important venue for information sharing, especially during disasters, but studies typically focus on individual tweets [1, 2]

02 **We focus on Twitter threads:** linked lists of tweets from a single author

03 We collected ~1000 coronavirus-related Twitter threads and annotated the ones with over 50 likes (163 threads).

04 This project helps:

- Organizations understand how they can share info on Twitter
- Researchers understand how info is spreading on Twitter

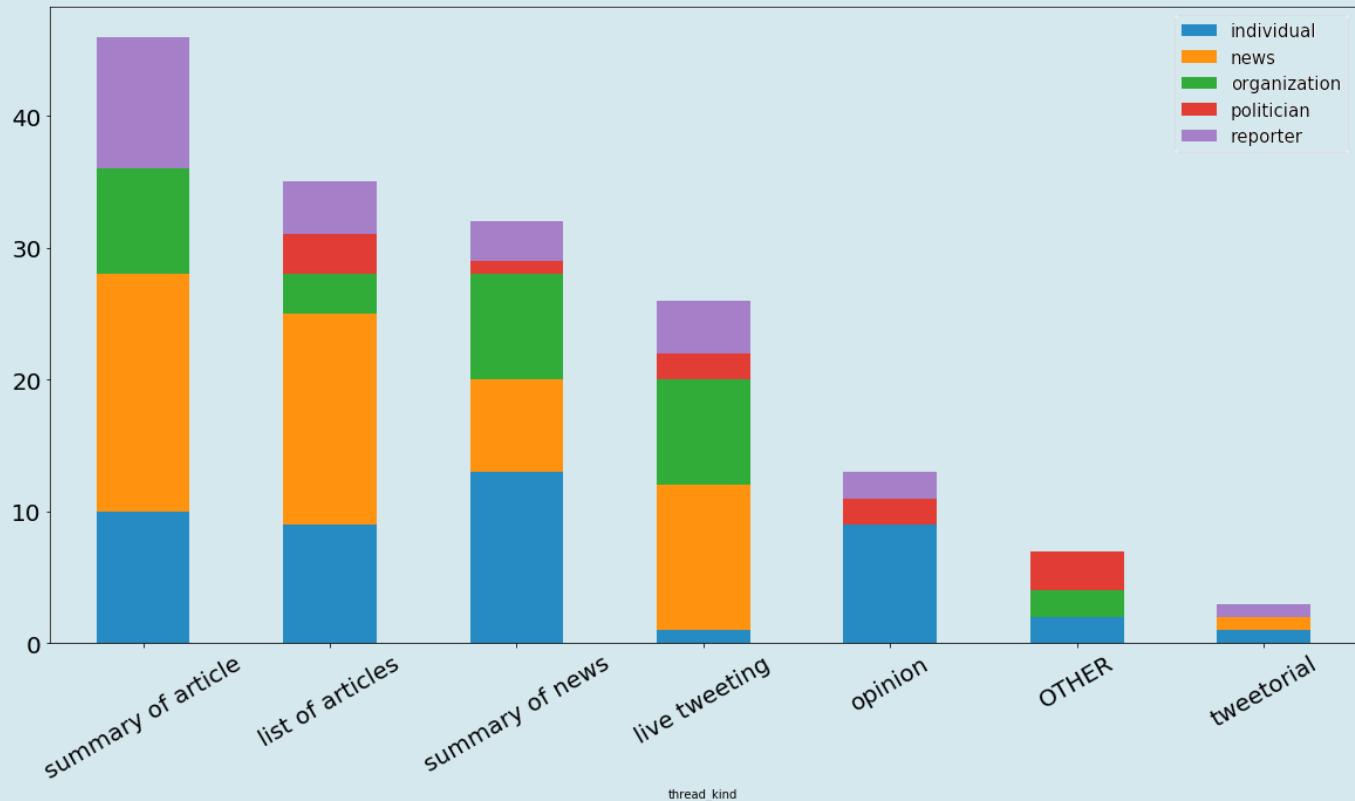
1. Stowe, Kevin, et al. "Developing and evaluating annotation procedures for Twitter data during hazard events." *Proceedings of the Joint Workshop on Linguistic Annotation, Multiword Expressions and Constructions (LAW-MWE-CxG-2018)*. 2018.
2. Anderson, T. Jennings, et al. "Far Far Away in Far Rockaway: Responses to Risks and Impacts during Hurricane Sandy through First-Person Social Media Narratives." *ISCRAM*. 2016.

## Data collection and annotation

- Collected tweets with 'coronavirus' or 'covid19' in the text using [twarc](#)
- On a daily basis search for threads: tweets that include 3+ replies from the same author
- **We collected ~1k Twitter threads from 4/9-4/19**
- **We annotated threads with 50+ likes (163 in total) according to:**
  - Locality (focus of content, e.g. USA, global, India)
  - Source (kind of author, e.g. reporter, politician, news organization)
  - Type (structure of content, e.g. summary of article, live tweeting)
- Additional Twitter data includes # retweets, # likes, user profile info



# Most Twitter threads are about external news



- Most threads summarize or refer to external news sources
- Most threads are tweeted by news organizations

## Next steps:

- How many readers read the external sources?
- Are there predictors of likes or retweets other than the author of the thread?

# Surprising uses of Twitter threads

**StatiSense** @StatiSense

WHAT IS THE IMPACT OF #COVID19 ON THIS INDUSTRY IN NIGERIA?

1. AGRICULTURE

SUBCATEGORY:  
 #1 Crop Production  
 #2 Livestock  
 #3 Forestry  
 #4 Fishing

**#StatiSense**

POSITIVE	29.2%
NEUTRAL	25.3%
<b>NEGATIVE</b>	<b>45.5%</b>

407 votes - Final results

7:49 AM · Apr 12, 2020 · Twitter for Android

9 Retweets 5 Likes

**StatiSense** @StatiSense · Apr 12  
 Replying to @StatiSense  
 WHAT IS THE IMPACT OF #COVID19 ON THIS SECTOR?

2. MINING AND QUARRYING

SUBCATEGORY:  
 #1 Crude Petroleum and Natural Gas

To gather data

This thread was a series of polls from a data science organization

**Barton** @Barton\_options

Coronavirus, mutations, and remdesivir

A story of Risk/reward in medicine and virology

Many people worry about COVID19 mutating as it spreads across the globe, for very good reasons:

**Will the antibodies of those recovered be effective against new infections?**

8:16 AM · Apr 12, 2020 · Twitter for iPad

30 Retweets 86 Likes

**Barton** @Barton\_options · Apr 12  
 Replying to @Barton\_options  
 Will mutations make vaccine useless? Drug resistance? More virulent? Etc. etc.

It turns out, coronavirus is special among RNA viruses with its unique proofreading protein. Google "coronavirus proofreading" and you can find a large number of recent research on that topic.

**Barton** @Barton\_options · Apr 12  
 as a result, coronavirus is relatively stable. Vaccines and antibodies are expected to be effective against it for a long time (2 year and above)

To explain a concept

This thread explains science behind antibodies as protection

**Presidency | South Africa** @PresidencyZA

#COVID19 | LOCKDOWN REGULATIONS - Categorisation of #Essential Goods #CoronaVirusSA #Day22ofLockdown

**SA LOCKDOWN REGULATIONS**

**CATEGORISATION OF ESSENTIAL GOODS:**

- 1. Food**
  - Any food product, including non-alcoholic beverages;
  - Animal feed; and
  - Chemicals, packaging and ancillary products used in the production of any food product.
- 2. Cleaning and Hygiene Products**
  - Toilet paper, sanitary pads, sanitary tampons, condoms;
  - Hand sanitiser, disinfectants, soap, alcohol for industrial use, household cleaning products, and personal protective equipment; and
  - Chemicals, packaging and ancillary products used in the production of any of the above.
  - Products for the care of hands and bodies.
  - Personal toiletries, including hygiene, body and face washes, roll-ons, deodorants, toothpaste.
- 3. Medical**
  - Medical and Hospital Supplies, medical equipment and personal protective equipment; and
  - Chemicals, packaging and ancillary products used in the production of any of the above.
- 4. Fuel, including coal, wood and gas.**
  - Basic goods, including airline and electricity and the withdrawal of cash.
- 5. Basic goods, including airline and electricity and the withdrawal of cash.**
  - Basic goods, including airline and electricity and the withdrawal of cash.
- 6. Hardware components and supplies required for essential repairs related to the provision of water, electricity or other essential services.**
  - Components for vehicles under-going emergency repairs where such vehicle is owned by or used by a person engaged in essential services work.
- 7. Components for vehicles under-going emergency repairs where such vehicle is owned by or used by a person engaged in essential services work.**
  - Components for vehicles under-going emergency repairs where such vehicle is owned by or used by a person engaged in essential services work.

WHATSAPP SUPPORT  
**0600 123 456**  
 EMERGENCY NUMBER  
**0800 029 999**  
 sacoronavirus.co.za

REPUBLIC OF SOUTH AFRICA

NDP 2030

7:11 AM · Apr 17, 2020 · Twitter for iPhone

123 Retweets 197 Likes

Official gov communication

This is a thread of announcements from South Africa's president

# Why were so few personal stories collected?

## 1. Method of data collection

- Often people will have a series of Tweets that are not explicitly a thread.
- Individuals are less likely to have many followers leading to fewer likes.

## 2. Format of individual tweets

- Character limit inhibits free-form writing.
- People may prefer to tweet personal stories on private accounts.

## To improve future data collection:

- Expand search to indirect terms (such as “pandemic”, “virus”, “quarantine”, “isolation”)
- Filter out tweets made by news organizations and focus a bigger part of search on individuals

Examples of personal stories not identified by our algorithm:

