

# Mining Social Media for COVID-19 Information

**Ashlynn Daughton, Dax Gerts, Courtney Shelley, Nidhi Parikh, Geoffrey Fairchild, Travis Pitts, Chrism Watson Ross, Nidia Vaquera**

# Objectives

- Mine social media for information that we can use to:
  1. Understand types of misinformation relevant to COVID-19 & how they are spreading
  2. Understand compliance with a variety of behaviors that are relevant for COVID-19 transmission

# Motivation



Image Credit: New York Magazine



Image Credit: Politico



Image Credit: Financial Times

Never said the virus wasn't real, but the Pandemic is a hoax in most states. Take this NY undercover nurse for example, stuff like this has been happening from the start of the Pandemic. Also we are testing more so the numbers will rise for a short period but go back down and we should also expect a rise as we reopen which is totally normal. We are dealing with a virus that has killed less than the flu. I study Population Health and can give you some good statistics on this. Not to mention the CDC and WHO have admitted to fudging the numbers, it's just a joke at this point <https://www.youtube.com/watch?v=UIDSkdeFomQ>



Perspectives on the Pandemic | The (Underco...  
youtube.com

Image Credit: Facebook

# Background - Misinformation

- Misinformation is a "claim of fact that is currently false due to lack of scientific evidence". (Kouzy et al. 2020)
- Conspiracy theories are a subset of misinformation
- People who believe conspiracy theories are often
  - Quite intelligent, but **lack scientific sophistication** to spot poor logic (Uscinski & Parent 2014, Ståhl & van Prooijen 2018)
  - Usually **lonely** and **isolated** (Gorman & Gorman 2017)
  - **Feel powerless by uncertainty** (Whitson & Galinsky 2008, Douglas et al. 2019)
  - **Need quick resolution** to the unknown. (Marchlewska et al. 2018)
- Attempts to correct misperceptions will often have a "**backfire effect**", making a person believe in the conspiracy theory even more firmly. (Lord et al. 1979)

# Data

- Twitter
  - Tweets with COVID-19 keywords (N ~ 540 million tweets, January 21 - Present) (Chen et al.)
- Reddit
  - Submissions — (N = 192 million)
  - Submissions & comments with COVID-19 keywords — (N = 1,359,966, includes 41,826 subreddits)
- COVID-19 Myths & Information Trustworthiness Scores (via NewsGuard — [www.newsguardtech.com](http://www.newsguardtech.com))

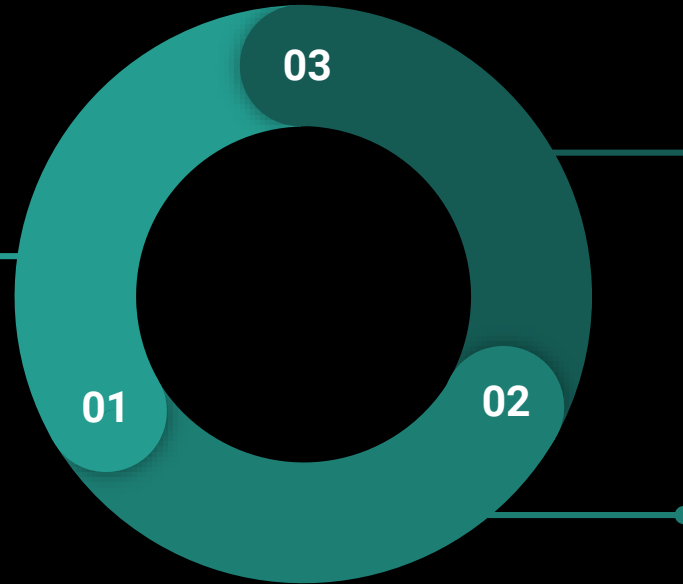
# Misinformation Methods

## Downstream analysis

- Geospatial, temporal & sentiment analysis
- Unsupervised learning (Dynamic Topic Modeling)

## Classification

- Label datasets
- Supervised machine learning



## Keyword filtering

- Filter for 13 top myths identified by NewsGuard.
- Exploratory analysis
  - temporal analysis
  - network analysis

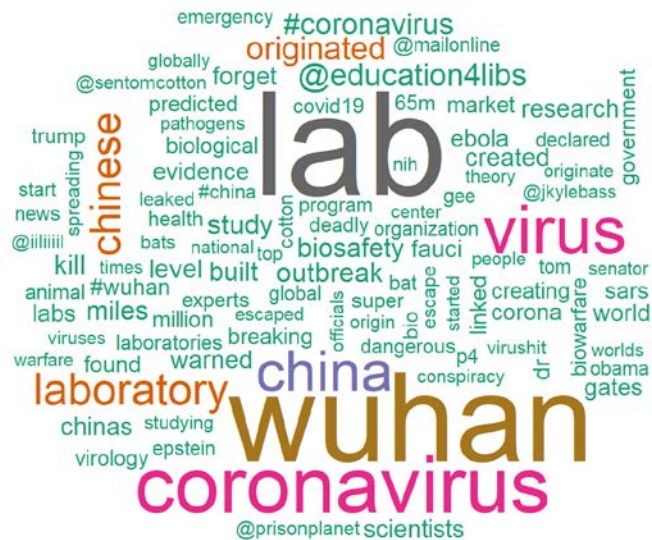
# Focus on 4 specific conspiracy theories

1. **“Origin story”** — COVID is bioweapon, or engineered in a laboratory
2. **5g and COVID**— COVID is actually the result of radiation from 5g, either through direct transmission or immune system impact
3. **Gates Foundation Created** — The Gates foundation funded a company which holds a patent on the coronavirus & is using this pandemic for their own personal financial gain
4. **Anti-vaccination/ Vaccine hesitancy** — The intersection of anti-vaccination or vaccine hesitant rhetoric and COVID. Includes people talking about why they won't get a COVID vaccine, and some overlap with the Gates foundation conspiracy (re: microchipping people for depopulation purposes)

# 2 Main Findings

## 1. History repeats itself

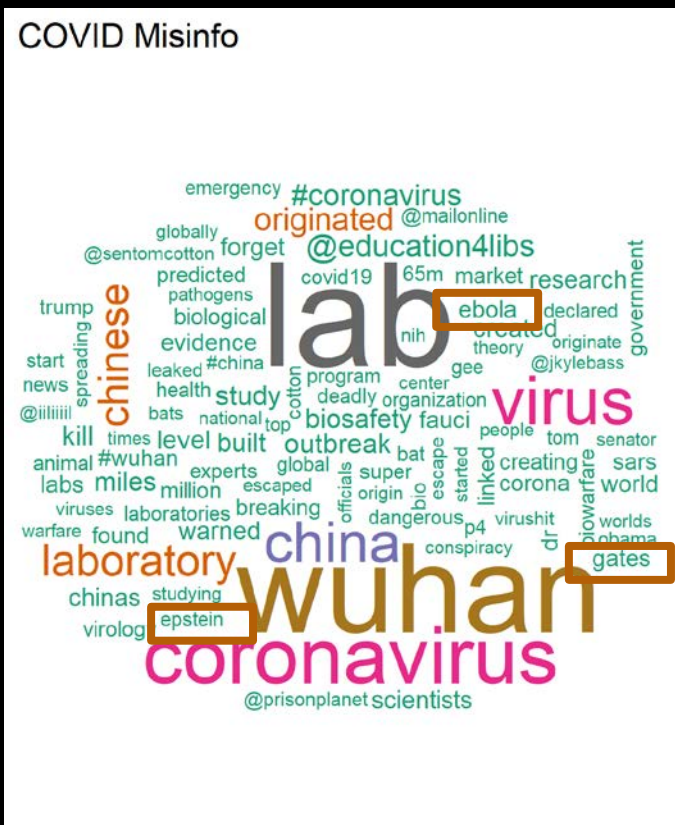
### COVID Misinfo





# 2 Main Findings

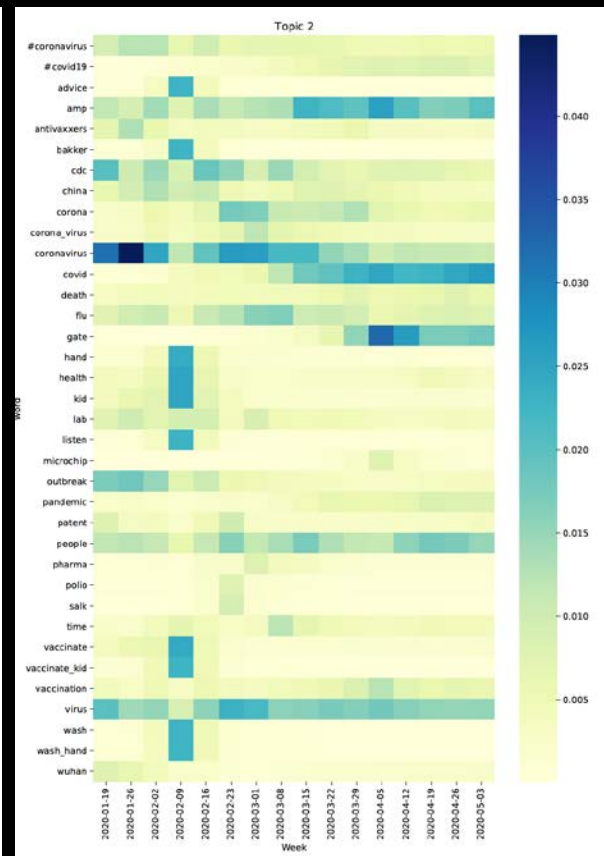
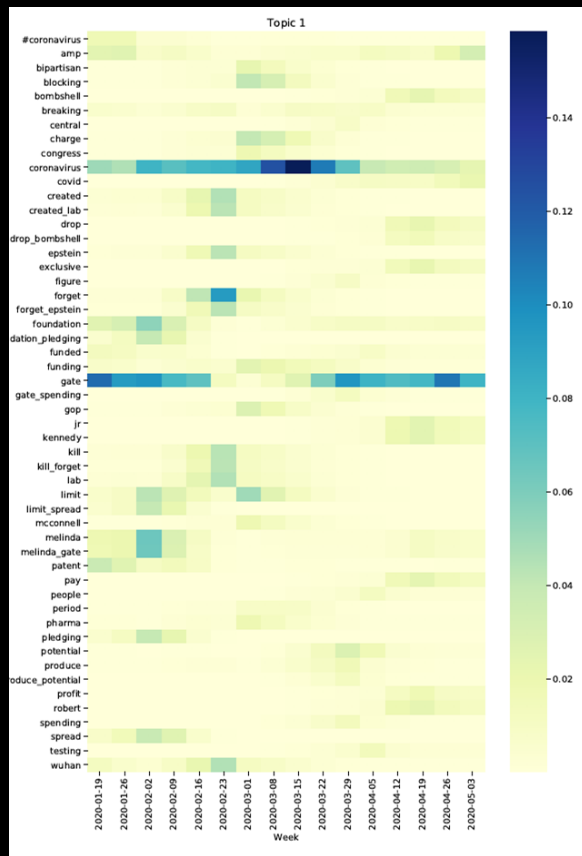
## 1. History repeats itself



# 2 Main Findings

1. History repeats itself

2. Conspiracy Theories Evolve



Questions?

EMAIL: [adaughton@lanl.gov](mailto:adaughton@lanl.gov)