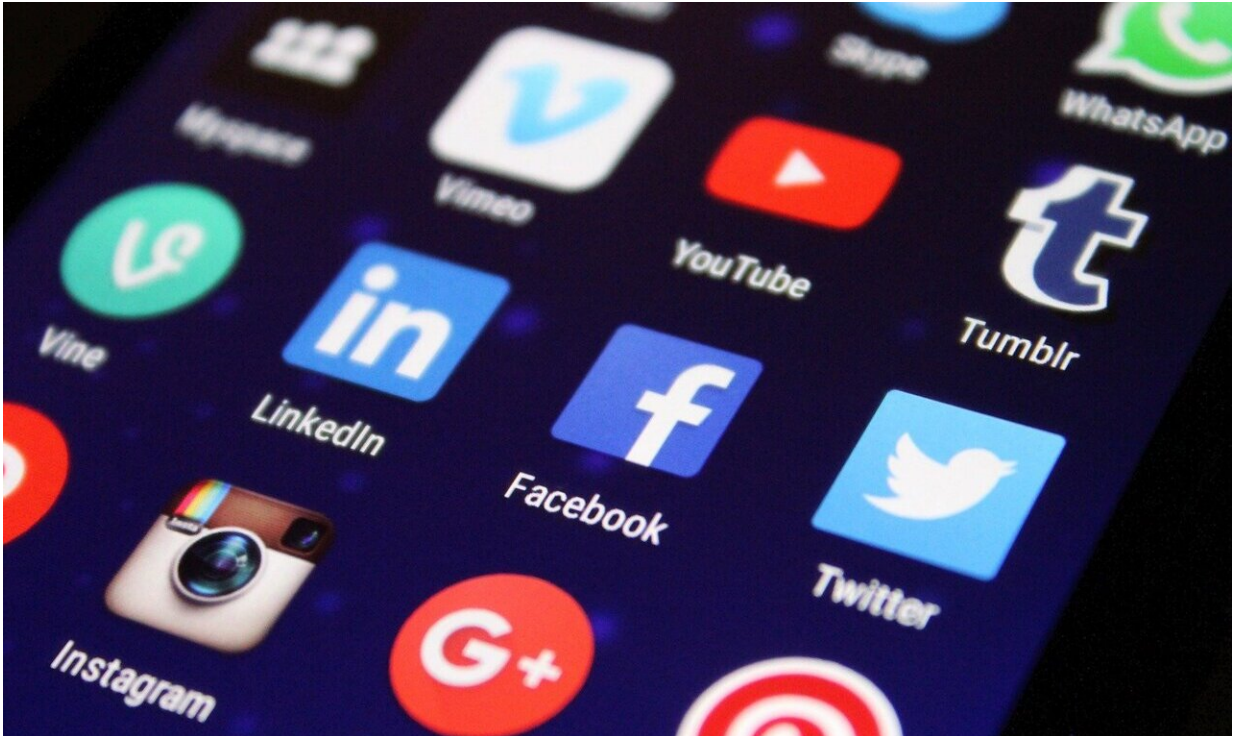


Using social media to map disease

October 22 2020, by Noelle Toumey Reetz



Credit: CC0 Public Domain

People around the world compose and send more than 500 million Tweets every day—and in the middle of a global pandemic, there's a lot to be learned from those geotags and hashtags. The massive amounts of data created by social media sites like Twitter is helping researchers gain powerful insights into the COVID-19 outbreak.. One of the researchers at the forefront of this work is Georgia State computer scientist Juan M.

Banda, who specializes in machine learning and natural language processing.

"Each time we pick up our smartphones and share a Tweet or update our status, we're providing information about ourselves in the form of data," says Banda. "All those bits and bytes provide a snapshot of our lives. On a macro scale they can provide insights into society and populations on a number of fronts."

Banda, assistant professor in the Department of Computer Science, began a project in March to collect and analyze Twitter data related to COVID-19. To date, his lab has compiled more than 700 million Tweets, which have yielded insights on the spread of misinformation and how human mobility has driven the pandemic's next move. He has also studied the reported symptoms of so-called long-haulers—patients who continue to suffer from long-term health problems as a result of SARS-CoV-2 infection.



Credit: Meg Buscema

The team has made the dataset publicly available as a resource for the global research community, and it's already been downloaded more than 30,000 times. The data have also been used in several multi-national studies.

Banda says his passion for [big data](#) came during [graduate school](#) at Montana State University, where he worked on the huge volume of data being generated by NASA's Solar Dynamics Observatory mission. That led to work in astroinformatics and a Ph.D. in computer science. It was at Stanford University, during a postdoctoral fellowship, that he began focusing on medical informatics. Earlier this year, he received an award through the Stanford Aging & Ethnogeriatrics Research Center to study racial bias in phenotyping algorithms for older adults.

Banda's research focuses on electronic phenotyping, or a way to use electronic health records to build models that can classify clinical conditions based on patient characteristics. He plans to determine the roles that racial and age bias play in building these algorithms, and then develop and modify software to achieve greater equity for older minority patients.

During the pandemic—as scientists have been racing for answers—international collaboration has been vital, and Banda has taken part in several multi-national research studies. One [study](#) published in *The Lancet Rheumatology* looked at patient side effects from the controversial drug hydroxychloroquine, which was touted as an effective treatment for COVID-19 by President Trump. The paper has helped the European Medicines Agency provide evidence-based guidance on the drug. [Another](#) study compared thousands of patients diagnosed with COVID-19 to patients previously diagnosed with influenza.

At Georgia State, Banda frequently collaborates with public health professor Gerardo Chowell, using social media as an epidemiology tool.

In 2019, they published a study that used Twitter data to track the spread of mosquito-borne infectious diseases following natural disasters.

The pair are building tools and frameworks to study individuals' movement patterns to learn more about how diseases spread.

"Extracting faithful signals from [social media](#) and behavioral data to generate real-time predictions of the course of epidemics and pandemics is an incredibly exciting research area," says Chowell. "Dr. Banda is well positioned to make significant advances in this direction."

Those advances continue to shed light on how deadly diseases affect our lives as the global pandemic stretches into the better part of a year.

"Social media provides a mostly unfiltered medium to access people's minds," says Banda. "With data from hundreds of thousands or even millions of people, it is one of the most powerful and diverse sources of information to understand so much around the pandemic and other major events like natural disasters."

More information: Banda et al., A large-scale COVID-19 Twitter chatter dataset for open scientific research - an international collaboration. (2020). [DOI: 10.5281/zenodo.4104587](https://doi.org/10.5281/zenodo.4104587)

Provided by Georgia State University

Citation: Using social media to map disease (2020, October 22) retrieved 8 May 2024 from <https://medicalxpress.com/news/2020-10-social-media-disease.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is

provided for information purposes only.