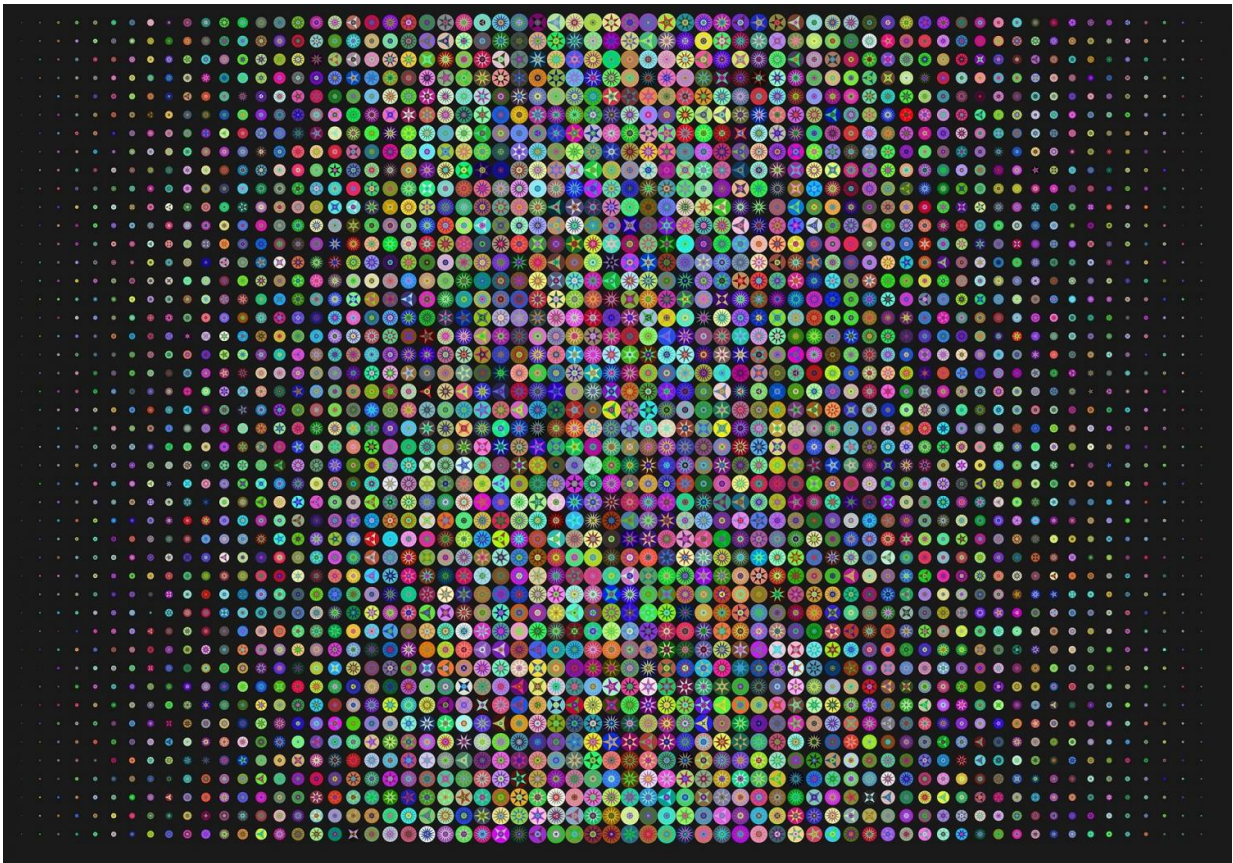# Algorithm tracker monitors Reddit rankings of COVID posts

April 15 2020, by Melanie Lefkowitz

Credit: CC0 Public Domain

On Reddit, which boasts 430 million monthly users, a trustworthy news article about California flattening its coronavirus curve was 77th in the day's recommended posts. Meanwhile, "Socially Distant Drinking

Games" was a top-ranked thread.

"During the pandemic, when we have limited time, there's lots of conflicting and competing information, and when we get suggestions from machines that say "Here are the top three things to read," that's probably what we're going to read," said J. Nathan Matias, assistant professor of communication in the College of Agriculture and Life Sciences. "And so paying attention to what these algorithms are promoting, and the quality of that information, is incredibly important."

Since 2016, Matias has tracked the algorithms on Reddit, a massive network of forums where people share content and news, and which claims to have more users than Twitter. As the coronavirus pandemic exploded, Matias began using the tool—called the COVID-19 Algo-Tracker—to monitor Reddit's virus-related posts and threads, both to inform people about the mechanisms behind the information they're receiving and to create a large, publicly available dataset for future research.

"The algorithm tracker was our attempt to take what we already have and create a public information resource about what kinds of pandemic information has been promoted by these powerful and influential algorithms that reach hundreds of millions of people," Matias said.

Developed by his Citizens and Technology Lab, the Algo-Tracker takes snapshots every two minutes of two of Reddit's key algorithms, including its default popularity ranking. The Algo-Tracker dashboard—updated every six hours—reviews posts that appear in the top 100 recommendations for both algorithms. Posts are considered coronavirus-related if they include any of 35 words or phrases—a periodically updated list currently including "self-isolate," "sanitizer," "Fauci," "flatten the curve," "chloroquine" and "herd immunity."

Examples of recent posts promoted by the algorithm are "Liberty University presses charges against journalists who covered campus being open during outbreak" and "A special effects artist made himself a mask for the coronavirus and I can't get over it."

In many cases, Matias said, the algorithm has promoted accurate scientific content, and some Reddit users have organized to try to provide reliable information. A new coronavirus community, created by users specifically to curate information about the pandemic, has been successful at having its posts promoted by Reddit's algorithms.

But human moderators are also finding themselves overwhelmed and overstretched by the volume of content—both reliable and questionable.

"The risk is that as people have conversations, as they try to cope with this incredible moment of stress and disruption, the communities may find themselves stretched to moderate and support reliable and accurate conversations about public health risks and what people can do about them," Matias said. "I think a lot of communities are finding themselves overwhelmed with a topic they didn't expect they'd be moderating and containing."

Different companies' algorithms work differently—and some large social media companies have said they've altered their algorithms to prioritize information from the Centers for Disease Control and Prevention and other reliable sources. But generally understanding how algorithms function and influence us has important implications for public health, Mattias said.

"You hear a rumor online—it's promoted to you by a social media algorithm—and then you go to Google to find out whether it's true. And the first thing you see on Google was chosen by an [algorithm](#)," he said. "Our social media networks are shaping what people are aware of during

the pandemic, and are at the heart of so much of what we know and believe and how we behave."

Provided by Cornell University