

# Reddit reveals peaks of public interest in COVID-19 topics

May 13 2020

---



Credit: CC0 Public Domain

Machine learning can be used to track surges in interest in health topics on popular online comment boards, like Reddit, according to a new study conducted during the COVID-19 outbreak by researchers in the

Perelman School of Medicine at the University of Pennsylvania (Penn Medicine). Such insight could help public health officials better understand and address public concerns and priorities, and stem the spread of misinformation. This study was published today in the *Journal of General Internal Medicine*.

"Public health priorities do not always align with community priorities, and the success of public health efforts often depends on having a plan to address community concerns," said Daniel Stokes, a research fellow with the Center for Emergency Care Policy and the Center for Digital Health at Penn Medicine. "Having a source like Reddit that is directly tied to people's thoughts could prove invaluable in crafting plans that meet people where they are."

The researchers chose to evaluate discussions on Reddit because it is one of the most popular sites on the internet, as well as being relatively unfiltered and up-to-date.

For example, researchers said real-time monitoring of Reddit could have allowed for a nimbler response during a surge of questions around whether it was safe to go outside in mid-March. The Centers for Disease Control and Prevention (CDC) did not issue official guidelines for safely enjoying parks and outdoor activities until early April. Stokes and his fellow researchers believe that if there had been more monitoring of online discussion activity, the guidance could have been issued closer to the peak of interest.

As a conduit directly to the thoughts of some people, Reddit is also valuable because it is the place where some of the "infodemic"—the plague of misinformation about COVID-19—has spread. Examples include one Reddit poster's belief that a natural remedy like licorice root might prevent COVID-19 infection, or another's thought that the virus was human engineered. Here, too, a quick, tailored response from public

health officials could lead to more fact-based and productive discourse.

To identify surges of interest in the public, the study's researchers collected nearly 95,000 posts from March 3 through March 31, 2020 on the most popular COVID-19 thread on Reddit, r/Coronavirus. They identified 50 different discussion topics through a [machine learning](#) technique of natural language processing. Then, 10 of those topics were determined to be most related to three areas of interest in the study: the response to public health measures, the sense of the pandemic's severity, and its impact on daily life.

By tracking how the popularity of these topics varied day-by-day, the team was able to demonstrate how areas of interest ebbed and flowed. For instance, hand-washing was found to peak early on, between March 3 and 6, while concern about personal finances was discussed roughly 50 percent more at the end of March as compared to the beginning. The analysis also showed that some topics popular at the start of the month remained top of mind, or had a comeback later in the month. Such was the case for mask-wearing.

"The CDC didn't make their recommendations on wearing masks in public until early April, so it is interesting to see that masks were being discussed a great deal prior to that recommendation," Stokes said.

"Perhaps it was a sign that many people were ready for these guidelines earlier."

Moving forward, the team will continue to track and analyze posts on this COVID-19-specific thread. Another effort from Penn's Center for Digital Health, led by Raina Merchant, MD, an associate professor of Emergency Medicine, has been to collect similar data through Twitter and map it across the United States.

"We are aiming to incorporate input from several digital sources that

would allow us to not just track the public's sentiment and perception of the virus, but also track, in real time, the emergence of new outbreaks," said Merchant, who is also the senior author of this *Journal of General Internal Medicine* study.

Stokes and Merchant hope insight like this will be heeded by [public health officials](#) in their effort to better combat the spread of misinformation that accompanied the COVID-19 outbreak.

"The success of our public health efforts depends on public buy-in," Stokes said. "Early comparisons to the flu on Reddit may have indicated a gap in public understanding of pandemic severity. Recognizing such gaps can be useful in developing targeted campaigns to close them."

**More information:** Daniel C. Stokes et al, Public Priorities and Concerns Regarding COVID-19 in an Online Discussion Forum: Longitudinal Topic Modeling, *Journal of General Internal Medicine* (2020). [DOI: 10.1007/s11606-020-05889-w](https://doi.org/10.1007/s11606-020-05889-w)

Provided by Perelman School of Medicine at the University of Pennsylvania

Citation: Reddit reveals peaks of public interest in COVID-19 topics (2020, May 13) retrieved 23 May 2024 from <https://medicalxpress.com/news/2020-05-reddit-reveals-peaks-covid-topics.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.