WILL KNIGHT   BUSINESS   FEB 8, 2020 7:00 AM

How AI Is Tracking the Coronavirus Outbreak

Machine-learning programs are analyzing websites, news reports, and social media posts for signs of symptoms, such as fever or breathing problems.

ILLUSTRATION: SAM WHITNEY; GETTY IMAGES

🗝 The AI Database →

APPLICATION: <u>PREDICTION</u>, <u>TEXT ANALYSIS</u>   END USER: <u>GOVERNMENT</u>, <u>RESEARCH</u>

SECTOR: <u>HEALTH CARE</u>, <u>PUBLISHING</u>, <u>SOCIAL MEDIA</u>   SOURCE DATA: <u>TEXT</u>, <u>GEOLOCATION</u>

TECHNOLOGY: <u>MACHINE LEARNING</u>, <u>NATURAL LANGUAGE PROCESSING</u>

WITH THE CORONAVIRUS growing more deadly in China, <u>artificial intelligence</u> researchers are applying <u>machine-learning</u> techniques to social media, web, and other data for subtle signs that the disease may be spreading elsewhere.

The new virus emerged in Wuhan, China, in December, triggering a <u>global health emergency</u>. It remains uncertain how deadly or contagious the virus is, and how widely it might have already spread. Infections and deaths continue to rise. More than 31,000 people have now contracted the disease in China, and 630 people have died, according to figures released by authorities there Friday.

<u>John Brownstein</u>, chief innovation officer at Harvard Medical School and an expert on mining social media information for health trends, is part of an international team using machine learning to comb through social media posts, news reports, data from official public health channels, and information supplied by doctors for warning signs the virus is taking hold in countries outside of China.

The program is looking for social media posts that mention specific symptoms, like respiratory problems and fever, from a geographic area where doctors have reported potential cases. Natural language processing is used to parse the text posted on social media, for example, to distinguish between someone discussing the news and someone complaining about how they feel. A <u>company called BlueDot</u> used a similar approach—minus the social media sources—to spot the coronavirus in late December, before Chinese authorities acknowledged the emergency.

"We are moving to surveillance efforts in the US," Brownstein says. It is critical to determine where the virus may surface if the authorities are to allocate resources and block its spread effectively. "We're trying to understand what's happening in the population at large," he says.

Ask WIRED

What is a coronavirus?

The rate of new infections has slowed slightly in recent days, from 3,900 new cases on Wednesday to 3,700 cases on Thursday to 3,200 cases on Friday, according to the World Health Organization. Yet it isn't clear if the spread is really slowing or if new infections are simply becoming more difficult to track.

So far, other countries have reported far fewer cases of coronavirus. But there is still widespread concern about the virus spreading. The US has imposed a travel ban on China even though experts question the effectiveness and ethics of such a move. Researchers at Johns Hopkins University have created a visualization of the virus's progress around the world based on official numbers and confirmed cases.

Health experts did not have access to such quantities of social, web, and mobile data when seeking to track previous outbreaks such as severe acute respiratory syndrome (SARS). But finding signs of the new virus in a vast soup of speculation, rumor, and posts about ordinary cold and flu symptoms is a formidable challenge. "The models have to be retrained to think about the terms people will use and the slightly different symptom set," Brownstein says.

Even so, the approach has proven capable of spotting a coronavirus needle in a haystack of big data. Brownstein says colleagues tracking Chinese social media and news sources were alerted to a cluster of reports about a flu-like outbreak on December 30. This was shared with the WHO, but it took time to confirm the seriousness of the situation.
Beyond identifying new cases, Brownstein says the technique could help experts learn how the virus behaves. It may be possible to determine the age, gender, and location of those most at risk more quickly than using official medical sources.

Alessandro Vespignani, a professor at Northeastern University who specializes in modeling contagion in large populations, says it will be particularly challenging to identify new instances of the coronavirus from social media posts, even using the most advanced AI tools, because its characteristics still aren't entirely clear. "It's

something new. We don't have historical data," Vespignani says. "There are very few cases in the US, and most of the activity is driven by the media, by people's curiosity."

But Vespignani believes that if the disease spreads more widely in the US, it should become easier to monitor its spread by applying machine learning to social media, news reports, and medical information. Combining AI with other techniques "could be very powerful," Vespignani says.

Crowdsourced information, collated by volunteers or via websites set up to offer information about the coronavirus, is also important to the effort. Brownstein is working with a Boston-based company, Buoy, that offers health advice to millions of people in the US online and through health provider portals. Buoy will offer advice for those who suspect they may have the coronavirus, feeding that to Brownstein and others as another data source.

Keep Reading

The latest on artificial intelligence, from machine learning to computer vision and more

An analysis of crowdsourced data from a Chinese physician community website, conducted by researchers at the National Institutes of Health, reveals a picture of delays in reporting new cases in Wuhan during the early stages of the pandemic. It also suggests that those younger than 15 years of age are more resilient.

Other signals may assist health officials in different countries prepare responses. Pings from mobile devices, along with flight and train itineraries, are helping epidemiologists build a picture of the spread of the virus and likely trajectory.

Andy Tatem, a professor at the UK's University of Southampton, and colleagues recently used anonymized historical data from smartphones, supplied by the Chinese search company Baidu, to model how the virus may have moved out of Wuhan in the days after it appeared.

Another group of researchers used data from Tencent, the Chinese company behind the popular Chinese app WeChat, to model the contagion. This suggests that the travel restrictions imposed by the Chinese authorities may have slowed the spread of the disease by a few days, providing critical time for countermeasures. Similar techniques could predict the spread through other countries should the contagion spread.

While it might be possible for the authorities to track individuals using the movement of their smartphones, Tatem says this is less useful than understanding broader trends and dynamics. And although it is unclear how widely the virus might yet travel, he says the biggest concern is that it could appear in countries with fewer health care resources to combat it. "Whether it can be contained in China, that's the question for the world right now," Tatem says.

More Great WIRED Stories

- A code-obsessed novelist builds a writing bot. The plot thickens
- Welcome to the era of supercharged lithium-silicon batteries
- Chris Evans goes to Washington
- The fractured future of browser privacy
- How to buy used gear on eBay—the smart, safe way
- 👁 The secret history of facial recognition. Plus, the latest news on AI
- 🏃 Want the best tools to get healthy? Check out our Gear team's picks for the best fitness trackers, running gear (including shoes and socks), and best headphones
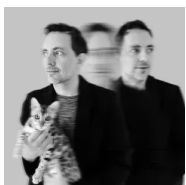
Will Knight is a senior writer for WIRED, covering artificial intelligence. He was previously a senior editor at *MIT Technology Review*, where he wrote about fundamental advances in AI and China's AI boom. Before that, he was an editor and writer at *New Scientist*. He studied anthropology and journalism in…
Read more

SENIOR WRITER   🐦

TOPICS   ARTIFICIAL INTELLIGENCE   MACHINE LEARNING   CORONAVIRUS   HEALTH   COVID-19

MORE FROM WIRED

Covid Exposure Apps Are Headed for a Mass Extinction Event

The US government will pull the plug on the servers powering the nation's Covid notifications on May 11. States aren't rushing to boot up replacements.

PARESH DAVE

Now That ChatGPT Is Plugged In, Things Could Get Weird

Letting the chatbot interact with the live internet will make it more useful—and more problematic, too.

WILL KNIGHT

Microsoft's 'Security Copilot' Unleashes ChatGPT on Breaches

The new tool aims to deliver the network insights and coordination that "AI" security systems have long promised.

LILY HAY NEWMAN

How ChatGPT—and Bots Like It—Can Spread Malware

Generative AI is a tool, which means it can be used by cybercriminals, too. Here's how to protect yourself.

DAVID NIELD

Workers Are Worried About Their Bosses Embracing AI

A survey by the Pew Research Center found that most employees expect hiring, firing, and workplace assessment to be transformed by algorithms.

WILL KNIGHT

The Hacking of ChatGPT Is Just Getting Started

Security researchers are jailbreaking large language models to get around safety rules. Things could get much worse.

MATT BURGESS

It's Way Too Easy to Get Google's Bard Chatbot to Lie

The company's policy bars use of the AI chatbot to "misinform." A study found that it readily spouted untruths on topics from Covid-19 to the war in Ukraine.

VITTORIA ELLIOTT

Amazon Is Joining the Generative AI Race

The ecommerce giant doesn't have a ChatGPT rival, but it wants to sell you the tools you need to build one.

WILL KNIGHT